

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 15:26:05 ; Search time 32 Seconds
(without alignments)
245.223 Million cell updates/sec

Title: US-10-768-874-2

Perfect score: 825

Sequence: 1 MSITQSFVLTALFQAASD.....NWGWMKGKHYFTVENCQE 152

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pdp.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pdp.*
- 5: /cgn2_6/ptodata/2/iaa/PCRUS COMB.pdp.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	825	100.0	152	3	US-09-036-355A-2
2	825	100.0	152	4	US-09-702-647A-2
3	610	73.9	109	3	US-09-036-355A-5
4	610	73.9	109	4	US-09-702-647A-5
5	130	15.8	22	3	US-09-036-355A-3
6	130	15.8	22	4	US-09-702-647A-3
7	88	10.7	712	4	US-09-059-584-51
8	87	10.5	15	3	US-09-036-355A-4
9	87	10.5	15	4	US-09-702-647A-4
10	84.5	10.2	1237	4	US-09-107-532A-4552
11	80.5	9.8	329	4	US-09-651-200-19
12	76	9.2	471	3	US-08-961-083-158
13	76	9.2	471	4	US-09-336-784-158
14	74	9.0	429	4	US-09-489-039A-12086
15	73.5	8.9	360	4	US-09-668-097A-32
16	72.5	8.8	911	1	US-08-487-890A-107
17	72.5	8.8	911	2	US-08-478-435-107
18	72.5	8.8	911	2	US-08-337-483-107
19	72.5	8.8	911	2	US-08-478-373-107
20	72.5	8.8	911	3	US-08-474-671-107
21	72.5	8.8	911	3	US-08-483-577A-107
22	72.5	8.8	911	3	US-08-897-438-107
23	72.5	8.8	911	4	US-08-637-654-107
24	72.5	8.8	911	4	US-08-649-518-107
25	72	8.7	616	4	US-09-608-790-1
26	72	8.7	943	4	US-09-397-885-5
27	72	8.7	943	4	US-09-969-362-5

Sequence 4219, Ap
Sequence 30, Appl
Sequence 30, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 20, Appl
Sequence 6366, Ap
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 1, Appl
Sequence 1, Appl

4 US-09-134-001C-4219
3 US-09-041-889-30
3 US-09-417-264-30
4 US-09-296-840A-4
2 US-08-793-824-2
4 US-09-651-200-20
4 US-09-543-681A-6366
1 US-08-487-890A-5
2 US-08-478-435-5
4 US-08-337-483-5
2 US-08-478-373-5
3 US-08-474-671-5
3 US-08-483-577A-5
3 US-08-897-438-5
4 US-08-637-654-5
4 US-08-649-518-5
4 US-09-397-885-1
4 US-09-969-362-1

ALIGNMENTS

RESULT 1
US-09-036-355A-2
; Sequence 2, Application US/09036355A
; Patent No. 6162785
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Swaim, Steven F.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM
; TITLE OF INVENTION: SALIVARY GLAND OF THE BLACK FLY
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSER: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,355A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5721-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-355A-2

Query Match 100.0%; Score 825; DB 3; Length 152;
Best Local Similarity 100.0%; Pred No. 4,1e-92;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSITQSFVLTALFQAASDNPFIADKRCIVISDGLVWHERKFGQFFYYVIMPKGTEY 60
DB 1 MSITQSFVLTALFQAASDNPFIADKRCIVISDGLVWHERKFGQFFYYVIMPKGTEY 60

QY 61 DDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAVREMDHFKFTADGTGKYDIS 120
DB 61 DDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAVREMDHFKFTADGTGKYDIS 120
QY 121 SKANGHPRSRGKNWGMKDGKHFTVENCQE 152
DB 121 SKANGHPRSRGKNWGMKDGKHFTVENCQE 152

RESULT 2

US-09-702-647A-2
; Sequence 2, Application US/09702647A
; Patent No. 6500420
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; TITLE OF INVENTION: Salivary Gland of the Black Fly
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/09/702,647A
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Simulium vittatum
US-09-702-647A-2

Query Match 100.0%; Score 825; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 4.1e-92;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSITQSFVLTALIFGAASDNPADKICIVISDGLVNHKPKGQEFFYVYVIMIPKGTGY 60
DB 1 MSITQSFVLTALIFGAASDNPADKICIVISDGLVNHKPKGQEFFYVYVIMIPKGTGY 60
QY 61 DDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAVREMDHFKFTADGTGKYDIS 120
DB 61 DDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAVREMDHFKFTADGTGKYDIS 120
QY 121 SKANGHPRSRGKNWGMKDGKHFTVENCQE 152
DB 121 SKANGHPRSRGKNWGMKDGKHFTVENCQE 152

RESULT 3

US-09-036-355A-5
; Sequence 5, Application US/09036355A
; Patent No. 6162785
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM
; TITLE OF INVENTION: SALIVARY GLAND OF THE BLACK FLY
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,355A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5721-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-036-355A-5

Query Match 73.9%; Score 610; DB 3; Length 109;
Best Local Similarity 100.0%; Pred. No. 3e-66;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 KPGQEFFYVYVIMIPKGTGYDDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAV 101
DB 1 KPGQEFFYVYVIMIPKGTGYDDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAV 60
QY 102 REMDHFKFTADGTGKYDISSKANGHPRSRGKNWGMKDGKHFTVENC 150
DB 61 REMDHFKFTADGTGKYDISSKANGHPRSRGKNWGMKDGKHFTVENC 109

RESULT 4

US-09-702-647A-5
; Sequence 5, Application US/09702647A
; Patent No. 6500420
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; TITLE OF INVENTION: Salivary Gland of the Black Fly
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/09/702,647A
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Simulium vittatum
US-09-702-647A-5

Query Match 73.9%; Score 610; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 3e-66;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 KPGQEFFYVYVIMIPKGTGYDDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAV 101
DB 1 KPGQEFFYVYVIMIPKGTGYDDQWILESGDHYKLNKPSGRYLVYGTDFYLTAGAAV 60
QY 102 REMDHFKFTADGTGKYDISSKANGHPRSRGKNWGMKDGKHFTVENC 150
DB 61 REMDHFKFTADGTGKYDISSKANGHPRSRGKNWGMKDGKHFTVENC 109

RESULT 5

US-09-036-355A-3
; Sequence 3, Application US/09036355A
; Patent No. 6162785
; GENERAL INFORMATION:

APPLICANT: Cupp, Mary S.
APPLICANT: Ribeiro, Jose M.C.
APPLICANT: Cupp, Eddie W.
APPLICANT: Swaim, Steven F.
TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM SALIVARY GLAND OF THE BLACK FLY
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
STREET: 3605 Glenwood Ave. Suite 310
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: 27622
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/036,355A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spruill, W. Murray
REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: 5721-4
TELEPHONE: 919 420 2202
TELEFAX: 919 881 3175
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-036-355A-3

Query Match 15.8%; Score 130; DB 3; length 22;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 GKNGVMKDGKHYFTVENCQE 152
Db 1 GKNGVMKDGKHYFTVENCQE 22

RESULT 6
US-09-702-647A-3
Sequence 3, Application US/09702647A
Patent No. 6500420
GENERAL INFORMATION:
APPLICANT: Cupp, Mary S.
APPLICANT: Ribeiro, Jose M.C.
APPLICANT: Cupp, Eddie W.
APPLICANT: Swaim, Steven F.
TITLE OF INVENTION: Recombinant Vasoactive Protein from Salivary Gland of the Black Fly
FILE REFERENCE: 35721/205407
CURRENT APPLICATION NUMBER: US/09/702,647A
CURRENT FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 22
TYPE: PRT
ORGANISM: Simulium vittatum
US-09-702-647A-3

Query Match 15.8%; Score 130; DB 4; length 22;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 GKNGVMKDGKHYFTVENCQE 152
Db 1 GKNGVMKDGKHYFTVENCQE 22

RESULT 7
US-09-059-584-51
Sequence 51, Application US/09059584
Patent No. 6440701
GENERAL INFORMATION:
APPLICANT: Myers, Lisa E
APPLICANT: Schryvers, Anthony B
APPLICANT: Harkness, Robin E
APPLICANT: Loosmore, Sheena M.
APPLICANT: Du, Run-Pan
APPLICANT: Yang, Yan-Ping
APPLICANT: Klein, Michel H
TITLE OF INVENTION: Transferrin Receptor Genes of Moraxella
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/059,584
FILING DATE: 14-APR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/778,570
FILING DATE: 03-JAN-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-794
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 712 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-059-584-51

Query Match 10.7%; Score 88; DB 4; length 712;
Best Local Similarity 27.8%; Pred. No. 0.09;
Matches 37; Conservative 10; Mismatches 44; Indels 42; Gaps 6;

Qy 38 MHERKQGPYVYVIMPKGTEDDQWILESVGGDHYKL-----KNKFSGR 84
Db 276 MTDKKGQSFSP-----ERRAGDEYSAMSSHEYSLLTDDKKNKPDNY 318
Qy 85 YLYVG-----TFDY---FLTAGAVERMDHFKFTADGTGKYDISSKANGHPRGRKNGV 136
Db 319 NDEYGHSSSEFTVDFSKSLTGLFSLNQDHHKGVTKTKYDINARTHGN-RFRGSATAI 377
Qy 137 MKDGE---KHYP 146
Db 378 NKDNESKAKHPFT 390

RESULT 8

US-09-036-355A-4
; Sequence 4, Application US/09036355A
; Patent No. 6162785
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM
; TITLE OF INVENTION: SALIVARY GLAND OF THE BLACK FLY
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,355A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5721-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-036-355A-4

Query Match 10.5%; Score 87; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 KPGQEFFPYVYMIPK 56
Db 1 KPGQEFFPYVYMIPK 15

RESULT 9
US-09-702-647A-4
; Sequence 4, Application US/09702647A
; Patent No. 6500420
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; TITLE OF INVENTION: Salivary Gland of the Black Fly
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/09/702, 647A
; CURRENT FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Simulium vittatum
US-09-702-647A-4

Query Match 10.5%; Score 87; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 KPGQEFFPYVYMIPK 56
Db 1 KPGQEFFPYVYMIPK 15

RESULT 10
US-09-107-532A-4552
; Sequence 4552, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4552:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1297 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1297
; SEQUENCE DESCRIPTION: SEQ ID NO: 4552:

US-09-107-532A-4552

Query Match 10.2%; Score 84.5; DB 4; Length 1297;
Best Local Similarity 28.8%; Pred. No. 0.57;
Matches 36; Conservative 14; Mismatches 36; Indels 39; Gaps 8;

QY 30 VISGDLVWHE-----RKPGQEFFPYV-YMI-----PKGTEVDOR--WIL-----ES 69
Db 920 VIVDGEKVIHPIGEPDSFVYKDSSEETNYAYKIESONPEGFKEPDENDTWILLYDNA 979
QY 70 VGGDHYKLKFKSGRYL-----VGTDFYELTAGAAVEMDHFKFTADGT 114
Db 980 IGTTLKQKOTGENWQLYTYTDEEDRLWVAFTAYNTFDY---REIRVRKLDEGHGLDGA 1036

115 GKXDI 119
1037 G-FDI 1040

Db

RESULT 11
US-09-651-200-19
; Sequence 19, Application US/09651200
; Patent No. 6429303
; GENERAL INFORMATION:
; APPLICANT: Green et al
; TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
; TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
; TITLE OF INVENTION: Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-562 (CURA-62)
; CURRENT APPLICATION NUMBER: US/09/651,200
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152383
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/172909
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/183578
; PRIOR FILING DATE: 2000-02-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 19
; LENGTH: 329
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-651-200-19

Query Match 9.8%; Score 80.5; DB 4; Length 329;
Best Local Similarity 27.1%; Pred. No. 0.24;
Matches 29; Conservative 13; Mismatches 38; Indels 27; Gaps 4;

QY 1 MSITQSFFVTLTAIGAA-----SDNPIADKCVISDGD-LVMH 39
Db 7 MELNLTFLVNTLLYGAASMKSQAYFNKGTGELPCHFTNSQNSLDELVVFWDQDKLVLY 66

QY 40 ERKPGQEPFYVYMTPKG-TEYDDQKWILSVGGDHKKNKFGRY 85
Db 67 ELYRGKENPQVHRKYKGTSTFDKNWTLRL-----HNIQIKDKGLY 108

RESULT 12
US-08-961-083-158
; Sequence 158, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE: OCT-30-1997
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 471 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 158:
US-08-961-083-158

Query Match 9.2%; Score 76; DB 4; Length 471;
Best Local Similarity 25.2%; Pred. No. 1.4;
Matches 30; Conservative 19; Mismatches 48; Indels 22; Gaps 6;

QY 34 GDLVWHERKPGQEPFYVYMTPKGTEYDDQKWILSVGGDHKKNKFGRYL 86
Db 156 GAMQWNEWLYDPAYSAYFYLKSDGT-YANQEW--QKVGKQWYFYFKKMGYMARNEWQG--- 209

QY 87 VYGTEDYFLTAGAAVREMDHFKFTADGTGKYDISSKANGHPRSRGKNVGMKDGKHYF 145
Db 210 -----NYLTGSGA---MATDEVIMDGTTRYIPAASGELKEKKDLNVGM-VHRDGRKRYFP 259

RESULT 13
US-09-536-784-158
; Sequence 158, Application US/09536784
; Patent No. 6573082
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/536,784
; FILING DATE: 30-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: OCT-30-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB340P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 158:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 471 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 158:
US-09-536-784-158

Query Match 9.2%; Score 76; DB 4; Length 471;
Best Local Similarity 25.2%; Pred. No. 1.4;
Matches 30; Conservative 19; Mismatches 48; Indels 22; Gaps 6;

Matches 30; Conservative 19; Mismatches 48; Indels 22; Gaps 6;
QY 34 GDLVHHRKPGQEPFYVYMIKPGCTEYDDORWILESVGGDHYKX-----NKFSGRYL 86
DB 156 GAMQONELYPDPAISAYFLKSDGT-YANQEW--QKVGKQWYFFKKWGYMARNEWQG--- 209
QY 87 VYGTDFYELTAGAAVREMDHFKFTADGTGKYDISKANGHPRSRCKNWCVMKDGSKHYF 145
DB 210 -----NYILTGSAA---NATDEVIMDGTFTYIPAASGELKEKKDLNVGW-VHRDGRKYFF 259

RESULT 14

US-09-489-039A-12086
; Sequence 12086, Application US/09489039A
; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12086
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12086

Query Match 9.0%; Score 74; DB 4; Length 429;
Best Local Similarity 31.7%; Pred. No. 2.2;

Matches 32; Conservative 16; Mismatches 37; Indels 16; Gaps 6;

QY 2 SITQSFVLTALFAGASDNPIADRKCIIVISDGLVHHRKPGQEPFYVYMIKPGTYD 61
DB 169 SVSSQF--LTALLMTA---PLAQDQTVIVIKGDLV---SKPYIDITLHL-MKTFGVEVD 218
QY 62 DQWILESVGGDHYKXKNKFSGRYLVG---TFDYFLTAGA 99
DB 219 NQSYQRFVVRG---KQYQSPGDYLVGSDASSYFLAAGA 256

RESULT 15

US-09-668-097A-32
; Sequence 32, Application US/09668097A
; Patent No. 6673988

; GENERAL INFORMATION:

; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Klein, Ted
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Plant Lipases
; FILE REFERENCE: B31401 US NA
; CURRENT APPLICATION NUMBER: US/09/668,097A
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/157309
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 32
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-09-668-097A-32

Query Match 8.9%; Score 73.5; DB 4; Length 360;
Best Local Similarity 22.9%; Pred. No. 1.9;

Matches 38; Conservative 25; Mismatches 52; Indels 51; Gaps 8;

QY 3 ITQSFVLT-----TLAIFGAAS-----DNPIADRKCIIVISDG--DLVHHR 41
DB 219 NQSYQRFVVRG---KQYQSPGDYLVGSDASSYFLAAGA 256

Db 153 LSKSLFILGERFGNDYNAAIFGKSLDEVYTVVPHI INKVTSGVETLIGLGAVDVVPVGV 212
QY 42 KPGQEPFY--VYMIKPGCTEYD-----DQWILESVGGDHYKXKNKFSGRY 85
DB 213 LPICGCFPLYLTLYGSSNQSDYDGDGLRRFNDLSGYHNRLLRQIGIG---RLRSKYAGVR 268
QY 86 LVYGTDFYELTAGAAVREMDHFKFT-----ADGTGKYDISKA 123
DB 269 LMYG--DFYTQVAENVRSPRSFGLDYGLTVCCGASGGQSYNNYNNKA 312

Search completed: September 9, 2004, 15:32:04
Job time : 33 secs

C;Genetics:
A;Gene: cbpb

Query Match 9.5%; Score 78; DB 2; Length 448;
Best Local Similarity 23.8%; Pred. No. 6.5;
Matches 41; Conservative 19; Mismatches 38; Indels 74; Gaps 11;
Qy 43 PGQFFPYVYMPKTEYDDQW-----ILESGDHY- 75
Db 217 PGEKVHYDQIL-----EKDGYKLSYTAINGSYRVQLEAVKNPLGNSVLSSTGGTHYF 271
Qy 76 KLNKFSGRVLYVGT-EDYELTAGAA-----VREMDHK-FTA-DGFGKY----- 117
Db 272 KIRSAIKTEPLVSAIVIDYTPGKVKHQDILEKDGKWLSTYTAINGRRVYIQLEGVTS 331
Qy 118 -----DISS-----KANG-----HPSRG-KNKGVMKDGKHYF 145
Db 332 QNYQNSGNISSYSGNSSTVCGWKINGSWYHFKSGSKSTGWLKDGSSWY 383

RESULT 3

S46122
SNF2 protein homolog YBR245c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YBR1633
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 08-Sep-2002
C;Accession: S46122
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S46122
A;Molecule type: DNA
A;Residues: 1-1143 <ALJ>
A;Cross-references: EMBL:Z36114; MIPS:YBR245c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD.ISW1
A;Cross-references: SGD:S0000449
A;Map position: 2R
C;Superfamily: chromatin remodelling complex ATPase chain ISW1; bromodomain homology
C;Keywords: nucleus; transmembrane protein
F;305-321/Domain: transmembrane #status predicted <TMM>

Query Match 9.5%; Score 78; DB 2; Length 1143;
Best Local Similarity 23.3%; Pred. No. 19;
Matches 37; Conservative 34; Mismatches 56; Indels 32; Gaps 9;
Qy 5 QSEFFVLTALFGAASNDPIADRK---CIVISDGLVMHRRKPGQBF-----PYVYMIPK 56
Db 49 ENLFKILIAIF-----LQEKKYDC-ISSGSIM-----TASEEYLENLEKPFQVGLPH 95
Qy 57 GFEYDDQRMVLSVGDHYKYLK---NKFSGRVLYVGTFTYFLTAGAAVREMDHKFTADG 113
Db 96 DPESNKKRYLLKDANGKFDLEGTWTFEHLISLSGLFKHTIESKAA--KDPKFRQVLDV 153
Qy 114 TQYDILSSKANGHPSRGKNGVMKDGKHYTVENCQE 152
Db 154 LEE-----NKANG--KGGKHQDVRRKRKTEHEDAEELLE 186

RESULT 4

S07146
Gamma-s-crystallin - common carp
N;Alternate names: beta-s-crystallin
C;Species: Cyprinus carpio (common carp)
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 13-Aug-1999
C;Accession: S07146
R;Chang, T.; Chang, W.C.
Biochim. Biophys. Acta 910, 89-92, 1987
A;Title: Cloning and sequencing of a carp beta-s-crystallin cDNA.
A;Reference number: S07146; MUID:88000706; PMID:3307924
A;Accession: S07146
A;Molecule type: mRNA

A;Residues: 1-174 <CHA>
A;Cross-references: EMBL:M26834; NID:g213051; PIDN:AAA49206.1; PID:g213052
C;Superfamily: beta-crystallin
C;Keywords: duplication

Query Match 9.3%; Score 77; DB 2; Length 174;
Best Local Similarity 26.7%; Pred. No. 2.8;
Matches 36; Conservative 18; Mismatches 41; Indels 40; Gaps 9;
Qy 29 IVISDGLVMHRRKPGQFFPYVYVMIKGTEDYDDQW-----ILESGDHYKYL 77
Db 36 IRVESGAWVIYER---PNEFGYQVLTGRGEYDQRMGLNDRCLSCMKIHVFSGSYKI 92
Qy 78 ----KNKFSGRVLYV-----GTFDFLTAGAAVREMDHKFTADGFGKVDISKANGH 126
Db 93 QLYDKGDTGQ--VYESTEDCFVVDRT-----REVHSCK-VLDGIWIF-----YEH 138
Qy 127 PRSRGKNWGMKGE 141
Db 139 PNYRGRQY-LLEKGE 152

RESULT 5

F98050
Lysozyme (EC 3.2.1.17) [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C;Accession: F98050
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: F98050
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-501 <KUB>
A;Cross-references: GB:A8007317; PIDN:AAL00235.1; PID:g15459086; GSPDB:GN00174
C;Genetics:
A;Gene: lytC
C;Keywords: glycosidase; hydrolase

Query Match 9.3%; Score 77; DB 2; Length 501;
Best Local Similarity 25.0%; Pred. No. 9.3;
Matches 31; Conservative 18; Mismatches 43; Indels 32; Gaps 8;
Qy 34 GDLVMHRRKPGQFFPYVYVMIKGTEDYDDQRMVLSVGDHYKYLK-----NKFSGRYL 86
Db 186 GAMIQNEWLYDPAYSAYFLKSDGT-YANQEW--QKVGKWIYFKKGMARNEWQGNV- 241
Qy 87 VGTDFYFLTAGAAVRE---MD--HFKFTADGFGKVDISKANGHPSRGKNGVMKGE 141
Db 242 -----YLTGSGMATDEVIDMGARYIFAASGELK-----EKKOLNVGM-VHRDCK 285
Qy 142 KHYF 145
Db 286 RYFF 289

RESULT 6

D97210
extracellular neutral metalloprotease, NPRE [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 02-Jun-2003
C;Accession: D97210
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: D97210

A;Accession: JCT869
A;Molecule type: DNA
A;Residues: 1-411 <ASA>
A;Cross-references: DDBJ:AB078775
A;Accession: PC7199
A;Molecule type: protein
A;Residues: 28-42195-111;152-170 <AS2>
C;Comment: This protein that is a fungal cell wall-degrading enzyme, hydrolyzes beta-1,4
C;Genetics:
A;Gene: bg1M

Query Match 9.3%; Score 76.5; DB 2; Length 411;
Best Local Similarity 23.0%; Pred. No. 8.3;
Matches 35; Conservative 23; Mismatches 59; Indels 35; Gaps 6;

Qy 7 FFLVTLAIFFG-----AASNPIADKRCI-----VTSGLDVLWHERKPGQEFF 48
Db 227 FILNLAVGGNWPNGSNNSTPFSSMLVDYRVRYQDTGSNVISDGYYTIASKSGKVMD 286
Qy 49 YVVYMIPKG-----TEY--DDQRWILESGGDHYKLKNKFSGRILYVGTFDYFLTAGAA 100
Db 287 VVDVSTARGAKIQWTNYVANNORFVESTGDGYIKLTAHSGVKLDVPSS--TSTGVQ 344
Qy 101 VREMD-----HFKEFADTGCKYDISSKANG 125
Db 345 LQOWDDNSNAQRWKIVDVGGGYIKLVSKVSG 376

RESULT 9
B97725
outer membrane protein omp1 [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C;Accession: B97725
R;Ogata, H.; Adic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; RC
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: B97725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-768 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL02740.1; PID:g15619252; GSPDB:GN00173
C;Genetics:
A;Gene: omp1
C;Superfamily: protective surface antigen D-15

Query Match 9.3%; Score 76.5; DB 2; Length 768;
Best Local Similarity 23.1%; Pred. No. 17;
Matches 30; Conservative 24; Mismatches 47; Indels 29; Gaps 7;

Qy 34 GDVLWHERKPGQEFFPVYVMIPKGYEYDQDWILLESVGGDHYKLKNKFSGRY---LVXGT 90
Db 563 GHITTYDQTDNKIVPKNGLYVSGTOEF-----AGVGGNNKYIKHEVDGKYKSFIHNK 615
Qy 91 PDYFLTA-----GAARVEMDHFKFTADGTGKYDISSKANG--HPRGRKKNWGVMKD 139
Db 616 LTKLSASGGNTLGKKIIRISDFNL-----GDYSLRGFANGGVGREKVTNEGL--G 668
Qy 140 GKHY-FTVE 148
Db 669 GERYYKFGTE 678

RESULT 10
S74926
sensory transduction system regulatory protein sll0649 - Synecocystis sp. (strain PCC 6
N;Alternate names: protein sll0649
C;Species: Synecocystis sp.
A;Variety: FCC 6803
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C;Accession: S74926
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miwaizima, N.

O. K.; Okumura, S.; Shimpoo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74926
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-245 <KAN>
A:Cross-references: EMBL:D90902; GB:AB001339; NID:gl652027; PIDN:BAAL6966.1; PID:gl65204
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: *phoP*
C:Superfamily: *ompR* protein; response regulator homology
A:Keywords: phosphoprotein; transcription regulation
F:14-125/Domain: response regulator homology <RRH>
F:62/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 9.2%; Score 75.5; DB 1; Length 245;
Best Local Similarity 30.7%; Pred. No. 5.8;
Matches 27; Conservative 10; Mismatches 20; Indels 31; Gaps 4;
QY 28 CIVISDG-----DLWMH-----RKPGQEPYVYVMPKGTG 59
DB 38 CIVTNDGTGLRIFQEQVDPDLVLDLMLPKLDGLEVCVTRIRQPGSKDPYILMLTAKGEE 97
QY 60 YDDQWILSVGGDHVYKLNKFSGRYLV 87
DB 98 ID-RILGLSTGADYLVK-PFSPRELV 122

RESULT 11
B64119
Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) - *Haemophilus influenzae* (strain R
C:Species: *Haemophilus influenzae*
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Jun-1999
C:Accession: B64119
R:Flerschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A>Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: B64119
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-444 <TIGR>
A:Cross-references: GB:U32815; GB:L42023; NID:gl574818; PIDN:AAC23006.1; PID:gl574822; T
C:Superfamily: glucose-1-phosphate adenylyltransferase
C:Keywords: glycogen/starch biosynthesis; nucleotidyltransferase
F:45/Binding site: fructose-1,6-bisphosphate (lys) #status predicted
F:120/Binding site: substrate (Tyr) #status predicted
F:202/Binding site: substrate (Tyr) #status predicted

Query Match 9.2%; Score 75.5; DB 2; Length 444;
Best Local Similarity 26.7%; Pred. No. 11;
Matches 27; Conservative 16; Mismatches 47; Indels 11; Gaps 3;
QY 45 QEPFYVYVMPKGTGTDQW---LLESVGGDHVYKLNKFSGRYLVYGTFFYFLTAGAAV 101
DB 99 QERGFEVDMFARQIDDSWYRGADAVYQNMALIKNHYRPKYILLAGDHIYKQDYSV 158
QY 102 REMDHFHTADGT-CKYDISSKANGHPRSRGKNGVMKDGK 141
DB 159 MLMHVNNGAKTVCIEV-----PRSEHFGVMVNE 192

RESULT 12
C95257
choline binding protein D [imported] - *Streptococcus pneumoniae* (strain TIGR4)
C:Species: *Streptococcus pneumoniae*
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C:Accession: C95257
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei
son, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison
A>Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95257
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-448 <KUR>
A:Cross-references: GB:AB005672; PIDN:AAK76252.1; PID:gl4973713; GSPDB:GN00164; TIGR:SP
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2201

Query Match 9.1%; Score 75; DB 2; Length 448;
Best Local Similarity 23.8%; Pred. No. 13;
Matches 41; Conservative 18; Mismatches 39; Indels 74; Gaps 11;
QY 43 PQGEPFYVYVMPKGTGTDQW-----ILESVGGDHV- 75
DB 217 PGEKVHYDQIL-----EKDGYKWLSTAYNGSYRYVQLEAVNKNPLGNSVLSSTGGTHYF 271
QY 76 KLNKFSGRYLVYGT-PDYFLTAGAA---VREMDHFK---FTA-DGTGKY----- 117
DB 272 KTRSAIKTEPLVSATVIDIYYPGKHYDQILEKDGKWLSTAYNGSRRIQLEGVTS 331
QY 118 -----DISS-----KANG-----HPRSNG-KNMGVMKDGSKHYF 145
DB 332 QNYQNGSNISYSGSHSSSTVGWKKINGSWYHFKSNGSKSTGWLKDGSSWY 383

RESULT 13
T48701
hypothetical protein IAG.50 [imported] - *Neurospora crassa*
C:Species: *Neurospora crassa*
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 05-May-2000
C:Accession: T48701
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24541
A:Accession: T48701
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-673 <SCH>
A:Cross-references: EMBL:AL353817; GSPDB:GN00112; NCSP:IA9.50
A:Experimental source: cosmid contig IAG; strain 74
C:Genetics:
A:Gene: NCSP:IA9.50
A:Map position: 2
A:Introns: 168/1; 316/2; 382/1; 626/3

Query Match 9.1%; Score 75; DB 2; Length 673;
Best Local Similarity 29.4%; Pred. No. 21;
Matches 32; Conservative 9; Mismatches 28; Indels 40; Gaps 7;
QY 35 DLVHMERKPGQEPFYVYVMPKGTGTDQWILSVGGDHVYKLNKFSGRYLVYGTFFYF 94
DB 572 DLSLPERKPG-----LQMGV--DCATWV--GVTV-----VVY----- 599
QY 95 LTAGAAVREMDHFKFTADGTGKYDISSKANGHPRSRGK-NMGVMKDGK 142
DB 600 -----ASRPDLFTFELDGECK--VKGVENGALREGEESNGGSGKEGR 641

RESULT 14
S00801
Na+/K+-exchanging ATPase (EC 3.6.3.9) alpha-3 chain - human
C:Species: *Homo sapiens* (man)
C>Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 19-Apr-2002
C:Accession: S00801; S04019; A27397; S02275

R;Ovchinnikov, Y.A.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Melkov, A.M.; Sverdlov, N.N.; Sverdlov, E.D.
 FEBS Lett. 233, 87-94, 1988
 A;Title: Family of human Na,K-ATPase genes. Structure of the gene for the catalytic subunit
 A;Reference number: S00801; MUID:88255304; PMID:2838329
 A;Accession: S00801
 A;Molecule type: DNA
 A;Residues: 1-1013 <OVC>
 A;Cross-references: EMBL:M37456
 R;Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Melkov, A.M.; Smirnov, N.N.; Ovchinnikov, Y.A.
 Dokl. Biochem. 297, 426-431, 1987
 A;Title: Family of human Na,K(+)ATPase genes. Structure of the gene of isoform alpha
 A;Reference number: S04019
 A;Accession: S04019
 A;Molecule type: DNA
 A;Residues: 1, 21H, 3-1013 <SVE1>
 A;Cross-references: EMBL:X12910; NID:G28963
 A;Note: The authors translated the codon TTC for residue 283 as Ser and TCT for residue 284 as Phe. This paper is a translation of the Russian paper published in Dokl. Akad. Nauk SSSR, 1987.
 R;Sverdlov, E.D.; Monastyrskaya, G.S.; Broude, N.E.; Ushkaryov, Y.A.; Allikmets, R.L.; Mäkinen, M.B.; Sverdlov, V.E.; Modyanov, N.N.; Ovchinnikov, Y.A.
 FEBS Lett. 217, 275-278, 1987
 A;Title: The family of human Na,K-ATPase genes. No less than five genes and/or pseudogenes
 A;Reference number: A27397; MUID:87247232; PMID:3036582
 A;Accession: A27397
 A;Molecule type: mRNA
 A;Residues: 243-434 <SVE2>
 A;Cross-references: GB:M27570
 C;Genetics:
 A;Gene: GDB:ATP1A3
 A;Cross-references: GDB:119713; OMIM:182350
 A;Map position: 19q13.2-19q13.2
 A;Introns: 2/3; 31/3; 51/3; 119/3; 157/3; 202/3; 242/1; 331/3; 398/1; 435/2; 479/3; 544/3
 C;Superfamily: Na/K-transporting ATPase alpha chain; ATPase nucleotide-binding domain
 C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium transp
 F;86-110/Domain: transmembrane #status predicted <TM1>
 F;120-139/Domain: transmembrane #status predicted <TM2>
 F;140-280/Domain: intracellular #status predicted <INT2>
 F;281-303/Domain: transmembrane #status predicted <TM3>
 F;310-338/Domain: transmembrane #status predicted <TM4>
 F;339-776/Domain: intracellular #status predicted <INT3>
 F;777-773/Domain: ATPase nucleotide-binding domain homology <ATN>
 F;777-800/Domain: transmembrane #status predicted <TM5>
 F;839-864/Domain: transmembrane #status predicted <TM6>
 F;865-942/Domain: intracellular #status predicted <INT4>
 F;943-968/Domain: transmembrane #status predicted <TM7>
 F;969-1013/Domain: extracellular #status predicted <EXT>
 F;366/Active site: Asp (aspartylphosphate intermediate) #status predicted
 F;498/Binding site: ATP (Lys) #status predicted
 F;707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match 9.1%; Score 75; DB 1; Length 1013;
 Best Local Similarity 21.5%; Pred. No. 33;
 Matches 48; Conservative 23; Mismatches 56; Indels 96; Gaps 12;
 QY 13 AIFGAASDN-PIADR-----KCIVISDGL-VNHER-KPQEPFY-----YVY 52
 Db 421 AVFKGGQDNIPVLRDVGADASESALLKCIELSSGSKVMRERKNKVAIPFNSNKNYQL 480
 QY 53 MIPKGTBYDDQRWILESVGGDH-----YKLNKFSGRYL----- 86
 Db 481 SIHETDPNDRYLLVMKGAPERILDRCSITLLOKQKPLDEEMKAFQNYLELGLGE 540
 QY 87 -VYGFDFYFL-----TAGAAVREMDHFKFTADG----- 113
 Db 541 RVLGFCFYLLPEEYQPGFAP-DCDDVNFITDNLFCVGLMSWIGFPRAAIVDAVCKRSA 599
 QY 114 -----TGKYDISKANGHRSRGNKWKCKEKFHTVEN 149
 Db 600 GIKVIMVGDHPITAKA-----IAKGVGIISGNE---TVED 633

RESULT 15

C24639
 Na/K-ATPase (EC 3.6.3.9) alpha-3 chain - rat
 N;Alternate names: Na/K-transporting ATPase alpha (III) chain
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 30-Jun-1988 #sequence revision 23-Apr-1993 #text_change 19-Apr-2002
 C;Accession: C24639; S00514; E27180; A60470
 R;Shull, G.B.; Greeb, J.; Lingrel, J.B.
 Biochemistry 25, 8125-8132, 1986
 A;Title: Molecular cloning of three distinct forms of the Na,K-ATPase alpha-subunit f
 A;Reference number: A90512; MUID:87128908; PMID:3028470
 A;Accession: C24639
 A;Molecule type: mRNA
 A;Residues: 1-1013 <SHU>
 A;Cross-references: EMBL:M14513; NID:G203030; PIDN:AAA40777.1; PID:G203031
 A;Note: In the authors' translation 405-Ser is shown after residue 409 and, consequentl
 J;Hara, Y.; Urayama, O.; Kawakami, K.; Nojima, H.; Nagamune, H.; Kojima, T.; Ohta, T.;
 J. Biochem. 102, 43-58, 1987
 A;Title: Primary structures of two types of alpha-subunit of rat brain Na,K(+)ATPase
 A;Reference number: S00460; MUID:88032933; PMID:2822682
 A;Accession: S00514
 A;Molecule type: mRNA
 A;Residues: 1-907, 'C', '909-1013 <HAR>
 R;Herrera, V.L.M.; Emanuel, J.R.; Ruiz-Opazo, N.; Levenson, R.; Nadai-Ginard, B.
 J. Cell Biol. 105, 1855-1865, 1987
 A;Title: Three differentially expressed Na,K-ATPase alpha subunit isoforms: structural
 A;Reference number: A92749; MUID:88033255; PMID:2822726
 A;Accession: B27180
 A;Molecule type: mRNA
 A;Residues: 1, NL, 4-103, 'R', 105-113, 'E', 115-127, 'G', 129-148, 'Q', 150-151, 'T', 153-165, 'D'
 A;Cross-references: EMBL:M28648; NID:G205633; PIDN:AAA41672.1; PID:G205634
 R;Hsu, Y.M.; Guidotti, G.
 Biochemistry 28, 569-573, 1989
 A;Title: Rat brain has the alpha3 form of the (Na,K)ATPase.
 A;Reference number: A60470; MUID:89229049; PMID:2540801
 A;Accession: A60470
 A;Molecule type: protein
 A;Residues: 117-132; 586-595, 'X', 597-601 <HSU>
 C;Comment: The alpha-3 form appears to be highly ouabain-inhibitable, as is alpha-2 but
 C;Genetics:
 A;Gene: NKAA3
 C;Superfamily: Na/K-transporting ATPase alpha chain; ATPase nucleotide-binding domain
 C;Keywords: ATP; heterodimer; hydrolase; ion transport; phosphoprotein; potassium trans;
 F;86-110/Domain: transmembrane #status predicted <TM1>
 F;120-139/Domain: transmembrane #status predicted <TM2>
 F;140-280/Domain: intracellular #status predicted <INT2>
 F;281-303/Domain: transmembrane #status predicted <TM3>
 F;310-338/Domain: transmembrane #status predicted <TM4>
 F;339-776/Domain: intracellular #status predicted <INT3>
 F;777-773/Domain: ATPase nucleotide-binding domain homology <ATN>
 F;777-800/Domain: transmembrane #status predicted <TM5>
 F;839-864/Domain: transmembrane #status predicted <TM6>
 F;865-942/Domain: intracellular #status predicted <INT4>
 F;943-968/Domain: transmembrane #status predicted <TM7>
 F;969-1013/Domain: extracellular #status predicted <EXT>
 F;366/Active site: Asp (aspartylphosphate intermediate) #status predicted
 F;498/Binding site: ATP (Lys) #status predicted
 F;707,711,716/Active site: Asp, Asp, Lys #status predicted

Query Match 9.1%; Score 75; DB 2; Length 1013;
 Best Local Similarity 21.5%; Pred. No. 33;
 Matches 48; Conservative 23; Mismatches 56; Indels 96; Gaps 12;
 QY 13 AIFGAASDN-PIADR-----KCIVISDGL-VNHER-KPQEPFY-----YVY 52
 Db 421 AVFKGGQDNIPVLRDVGADASESALLKCIELSSGSKVMRERKNKVAIPFNSNKNYQL 480
 QY 53 MIPKGTBYDDQRWILESVGGDH-----YKLNKFSGRYL----- 86
 Db 481 SIHETDPNDRYLLVMKGAPERILDRCSITLLOKQKPLDEEMKAFQNYLELGLGE 540

```
Qy 87 -VYGFDFYFL-----TAGAAVREMDHFKFTADG----- 113
Db 541 RVLGFCHYILPEEQPKGFAP-DCDDVNFTTDNLCFVGLMSMIDPPRAAVPDVAGKCRSA 599
Qy 114 -----TKKYDIDSSKANGHPRSRGKNWGMKDGKHYFTVEN 149
Db 600 GIKVIMVTGDHPITAKA-----IAKGVIISEGNE---TVED 633
```

Search completed: September 9, 2004, 15:29:21
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 15:15:29; Search time 23 Seconds
(without alignments)
344.116 Million cell updates/sec

Title: US-10-768-874-2
Perfect score: 825
Sequence: 1 MSITQSFVLTALFAGASD.....NWGVKDGKHFTVENCQE 152

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	99.5	12.1	637	SCAA_CHICK	Q92075 gallus gall
2	82	9.9	1597	GTF1_STRDO	P11001 streptococ
3	78	9.5	1129	YB95_YEAST	P38144 saccharomyc
4	77	9.3	173	CRBS_CYPAC	P10112 cyprinus ca
5	76.5	9.3	640	SCAA_RABIT	O97741 crytologus
6	75.5	9.2	437	GLGC_HAENI	P43796 haemophilus
7	75	9.1	1013	ALIA3_HUMAN	P13637 homo sapien
8	75	9.1	1013	ALIA3_RAT	P06687 rattus norv
9	73.5	8.9	613	PEPF MYCFU	Q98QP0 mycoplasma
10	73	8.8	276	MYPR_CHICK	P23289 gallus gall
11	73	8.8	276	MYPR_PORCU	P47790 poephila gu
12	72.5	8.8	164	P825_ARATH	O27711 arabidopsis
13	72	8.7	616	ECAR_ECHCA	Q27711 arabidopsis
14	72	8.7	742	NSF_ARATH	Q90495 echis carin
15	71.5	8.7	358	SFR1_ARATH	Q90495 arabidopsis
16	71.5	8.7	358	RFL1_STAEP	O22315 arabidopsis
17	71.5	8.7	640	APRN_ENTHI	O8CNI7 staphylococ
18	71.5	8.7	656	ACSA_RHOCA	P20301 entameba h
19	71	8.6	362	OMPD_SALTY	O88040 rhodobacter
20	70.5	8.5	367	OMPC_ECOLI	P37592 salmonella
21	70.5	8.5	473	6FGD_BUCAP	P06996 escherichia
22	70.5	8.5	1139	INAI_CAEEL	Q92H99 buchnera ap
23	70	8.5	330	INTH3_HAENI	Q03600 caenorhabdi
24	70	8.5	441	XYLA_RHILO	P20589 haemophilus
25	69.5	8.4	275	LM22_VIBPA	Q98CR8 rhizobium l
26	69.5	8.4	603	GLMS_THETH	Q56714 vibrio para
27	69.5	8.4	778	RHP9_SCHPO	Q56213 t glucosami
28	68.5	8.3	185	COMI_DICDI	P87074 schizosacch
29	68.5	8.3	698	SCAA_RAT	Q03380 dictyosteli
30	68.5	8.3	711	TBBI_NEIMB	P37089 rattus norv
31	68.5	8.3	910	SYL_NEIMA	Q09057 neisseria m
32	68	8.2	480	DISA_TRIGA	Q9JW39 neisseria m
33	68	8.2	597	GLMS_PYRAB	P15503 trimeresuri
					Q9V249 p glucosami

ALIGNMENTS

RESULT 1

ID	SCAA_CHICK	STANDARD;	PRT;	637 AA.
AC	Q92075; P70095; Q98941;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Amiloride-sensitive sodium channel alpha-subunit (Epithelial Na+ channel alpha subunit) (Alpha ENAC) (Nonvoltage-gated sodium channel 1 alpha subunit) (SCNEA) (Alpha NaCh).			
DE	Amiloride-sensitive sodium channel alpha-subunit (Epithelial Na+ channel alpha subunit) (Alpha ENAC) (Nonvoltage-gated sodium channel 1 alpha subunit) (SCNEA) (Alpha NaCh).			
GN	SCNN1A OR ENAC.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
FN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).			
RC	STRAIN=Isa brown; TISSUE=Cochlea;			
RX	MEDLINE=97157073; PubMed=9003454;			
RA	Killick R., Richardson G.;			
RT	"Isolation of chicken alpha ENAC splice variants from a cochlear cDNA library."			
RL	Biochim. Biophys. Acta 1350:33-37(1997).			
FN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Intestine;			
RA	Goldstein O., Asher C., Garty H.;			
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: Sodium permeable non-voltage-sensitive ion channel inhibited by the diuretic amiloride. Mediates the electrodiffusion of the luminal sodium (and water, which follows osmotically) through the apical membrane of epithelial cells. Controls the reabsorption of sodium in kidney, colon, lung and sweat glands.			
CC	Also plays a role in taste perception.			
CC	-1- SUBUNIT: Heterotetramer of two alpha, one beta and one gamma subunit. A delta subunit can replace the alpha subunit (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=2;			
CC	Name=Long;			
CC	Isoids=Q92075-1; Sequence=Displayed;			
CC	Name=Short;			
CC	Isoids=Q92075-2; Sequence=VSP_006196, VSP_006197;			
CC	-1- TISSUE SPECIFICITY: The long isoform has been found in cochlea, colon, and cartilage. The short isoform is only found in cochlea.			
CC	-1- PTM: Ubiquitinated; this targets individual subunits for proteasome-mediated degradation (By similarity).			
CC	-1- SIMILARITY: Belongs to the amiloride-sensitive sodium channel family.			
CC	-1- CAUTION: It is uncertain whether Met-1 or Met-15 is the initiator.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its			

Q24298 drosophila
Q959n1 rhizobium m
P41343 mesembryant
Q8v116 bruceella me
Q8CV38 bruceella su
P41581 citrobacter
P41583 citrobacter
Q17967 caenorhabdi
O15375 homo sapien
Q27874 methanobact
Q9Jxt2 neisseria m
P03306 f genome po

34 68 8.2 1507 1 CADE_DROME
35 67.5 8.2 340 1 LPSE_RHIME
36 67.5 8.2 365 1 FENR_MESCR
37 67.5 8.2 398 1 TRMU_BRUME
38 67.5 8.2 398 1 TRMU_BRUSU
39 67.5 8.2 445 1 6PGD_CITAM
40 67.5 8.2 445 1 6PGD_CITFR
41 67.5 8.2 485 1 PDII_CAEEL
42 67.5 8.2 505 1 MOT6_HUMAN
43 67.5 8.2 565 1 SYG_METTH
44 67.5 8.2 876 1 SYL_NEIMB
45 67.5 8.2 2333 1 POLG_FMDV1


```

SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Aljinovic G., Pohl F.M., Pohl T.M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
[2]
RN REVISIONS.
RA Aljinovic G., Pohl F.M., Pohl T.M.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear [Potential].
CC -1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family. SNF2L
CC subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z36114; CAA95208.1; --
CC PIR; S46122; S46122.
CC GerMOnline; 138788; --
CC SGD; S0000449; ISW1.
CC GO; GO:0005634; C:nucleus; IPI.
CC GO; GO:0006338; P:chromatin modeling; IGI.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR001650; Helicase C.
CC InterPro; IPR001005; Myb_DNA_Binding.
CC Pfam; PF00271; helicase_C; 1.
CC Pfam; PF00176; SNF2_N; 1.
CC SMART; SM00487; DEXDC; 1.
CC SMART; SM00490; HELICc; 1.
CC SMART; SM00717; SANT; 2.
CC PROSITE; PS00690; DEAD_ATP_HELICASE; FALSE NEG.
CC Hypothetical protein; Hydrolase; Helicase; Nuclear protein;
CC ATP-binding; DNA-binding.
CC NP_BIND 221 228 ATP (BY SIMILARITY).
CC SITE 324 327 DEAD BOX.
CC SEQUENCE 1129 AA; 131101 MW; 2CF86BAF553E6695 CRC64;
Query Match 9.5%; Score 78; DB 1; Length 1129;
Best Local Similarity 23.3%; Pred. No. 8.2;
Matches 37; Conservative 34; Mismatches 56; Indels 32; Gaps 9;
QY 5 QSGFVLTALFGAASDNPADRK--CIVISQDGLVMHKKPGQEE-----PYIVVMIPK 56
Dbb 35 ENLFKILIAIF-----LQEKKKYDC--ISSGSIN-----TASERYLENLKKPFQVGLPPH 81
QY 57 GTEYDDQRWILSVGGDHYKJK--NKFSGRYLVYGTDFYFLTAGAAVREMDHFKFTADG 113
Dbb 82 DPESNKKRYLLKDANGKKPDLEGTGTRFPHLLSLGLFKHFIESKAA--KDKFRQVLVD 139
QY 114 TKGVIYSSKANGHPRSRGKNGWGMKDGEXHYFTVNCQE 152
Dbb 140 LEE---NKANG--KKGKHQDVRRRKTSHEEDAELLKE 172
RESULT 4
CRBS_CYPCA
IID CRBS_CYPCA STANDARD; PRT; 173 AA.
AC P10112;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Beta crystallin S (Gamma crystallin S).
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]

```

```

SEQUENCE FROM N.A.
TISSUE=Lens;
MEDLINE=88000706; PubMed=3307924;
Chang T., Chang W.-C.;
"Cloning and sequencing of a carp beta s-crystallin cDNA.";
Biochim. Biophys. Acta 910:89-92(1987).
RL FUNCTION: Crystallins are the dominant structural components of
the vertebrate eye lens.
CC -!- DOMAIN: Has a two-domain beta-structure, folded into four very
similar Greek key motifs.
CC -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M26834; AAA49206.1; --
DR PIR; S07145; S07146.
DR HSRP; P06504; IATH.
DR InterPro; IPR001064; Crystallin.
DR Pfam; PF00030; Crystall; 2.
DR PRINTS; PR01367; BGCRYSTALLIN.
DR SMART; SM00247; XTALbg; 2.
DR PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW Eye lens protein; Repeat.
KW INIT MET
FT 0
FT DOMAIN 1 39 BETA/GAMMA CRYSTALLIN 'GREEK KEY' 1.
FT DOMAIN 40 82 BETA/GAMMA CRYSTALLIN 'GREEK KEY' 2.
FT DOMAIN 83 88 CONNECTING PEPTIDE
FT DOMAIN 89 129 BETA/GAMMA CRYSTALLIN 'GREEK KEY' 3.
FT DOMAIN 130 172 BETA/GAMMA CRYSTALLIN 'GREEK KEY' 4.
SQ SEQUENCE 173 AA; 20697 MW; 2E5DBA95453D2897 CRC64;

Query Match 9.3%; Score 77; DB 1; Length 173;
Best Local Similarity 26.7%; Pred. No. 1.3;
Matches 36; Conservative 18; Mismatches 41; Indels 40; Gaps 9;

QY 29 IVISDGLVHMRKFGQFPFYVYVIMPRGTEYDQW-----ILSVGDDHYKL 77
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 35 IRVESGAMVYER---PNFMGYQYVLTGRGYPDIQRWGLNDRLCSCRMHIFVSGSEYKI 91
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 78 ----KVKSGRYLVY-----GTDFYFLTAGAAVREMDHMFKFTADGTGKYDIDSSKANGH 126
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 QLYDKGDFTGQ--VYESTEDCFSVVDREFT-----REVHSCK-VLDGIWIF-----YEH 137
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 127 PRSRGKNGVMKDG 141
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 138 PNYGRQY-LLEKGE 151
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
SCAA RABIT
ID SCAA RABIT STANDARD; PRT; 640 AA.
AC O97741; Q9N133;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Amiloride-sensitive sodium channel alpha-subunit (Epithelial Na+
channel alpha subunit) (Alpha ENaC) (Nonvoltage-gated sodium channel 1
alpha subunit) (SCN5A) (Alpha NaCh).
GN SCN5A.
OS Oryzotolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Kudlacek O., Weisz E., Wiener H., Plass H.;

```

140	LEE	---	NKANG	--	KGKGKHQD	VRRRKTE	THEEEDAE	LLKE	172
:	:		:		:	:	:	:	:

RESULT 4
CRBS CYPCA
ID - CRBS CYPCA
STANDARD: PPT; 173 AA.

AC	
DT	FI0112;
DD	01-WAR-1989 (Rel. 10, Created)
DT	01-APR-1990 (Rel. 14, last sequence update)
DT	10-OCT-2003 (Rel. 42, last annotation update)
DE	Beta crystallin S (Gamma crystallin S).
DS	Cyprinus carpio (Common carp).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; P
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysii
OC	Cyprinidae; Cyprinus.
OX	NCBI_TaxID=7962;
RN	[1]
	RN [1]

RT "The rabbit epithelial sodium channel.":
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 81-435 FROM N.A.
 RA Velazquez H., Silva T.C., Andujar E., Jaffer A., Ortiz D.;
 RT "The rabbit DCT does not express amiloride sensitive sodium channel.":
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Sodium permeable non-voltage-sensitive ion channel
 CC inhibited by the diuretic amiloride. Mediates the electrification
 CC of the luminal sodium (and water, which follows osmotically)
 CC through the apical membrane of epithelial cells. Controls the
 CC reabsorption of sodium in kidney, colon, lung and sweat glands.
 CC Also plays a role in taste perception.
 CC -!- SUBUNIT: Heterotrimer of two alpha, one beta and one gamma
 CC subunit. A delta subunit can replace the alpha subunit. Interacts
 CC with the WW domains of NEDD4, WWP1 and WWP2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- PTM: Ubiquitinated; this targets individual subunits for
 CC proteasome-mediated degradation (By similarity).
 CC -!- SIMILARITY: Belongs to the amiloride-sensitive sodium channel
 CC family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ132108; CAA10571.1; -;
 CC EMBL; AF229025; AAF43680.1; -;
 CC InterPro; IPR004724; Enac.
 CC Dr Pfam; PF00858; ASC, 1.
 CC Dr PRINTS; PR01078; AMINACHANNEL.
 CC Dr TIGRFS; TIGR00859; ENAC, 1.
 CC Dr PROSITE; PS01206; ASC, 1.
 CC Ion transport; Sodium transport; Ionic channel; Transmembrane;
 CC Glycoprotein; Ubl conjugation; Sodium channel.
 CC KW DOMAIN 1 65 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 66 86 POTENTIAL.
 CC FT DOMAIN 87 539 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 540 560 POTENTIAL.
 CC FT TRANSMEM 561 640 POTENTIAL.
 CC FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 211 211 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CONFLICT 222 222 S -> P (IN REF. 2).
 CC FT CONFLICT 264 265 AF -> TS (IN REF. 2).
 CC FT CONFLICT 372 373 TQ -> AE (IN REF. 2).
 CC SEQUENCE 640 AA; 72734 MW; 17C84C100F69E133 CRC64;
 CC
 CC Query Match 9.3%; Score 76.5; DB 1; Length 640;
 CC Best Local Similarity 30.5%; Pred. No. 6.2;
 CC Matches 29; Conservative 10; Mismatches 31; Indels 25; Gaps 5;
 CC
 CC QY 49 YVYVMPKGYEYDQR---WILESGVGHYKLNKFSGRYLYVGTG-----DYF 94
 CC Db 410 YIYFPEGEYCYDKRHSN-----GYCYKLDQAFSSRL--GCTKCKPCTVNYE 462
 CC
 CC QY 95 LTAG----AAVREMDHFKFTADGTGKYDISSKANG 125
 CC Db 463 LSAGYSEWPSVTSQDWVFMQLSQNNYVTSNKGK 497
 CC
 CC RESULT 6
 CC GLGC_HABIN STANDARD; PRT; 437 AA.
 CC ID GLGC_HABIN

P43796;
 01-NOV-1995 (Rel. 32, Created)
 10-OCT-2003 (Rel. 42, Last sequence update)
 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27) (ADP-glucose
 DE synthase) (ADP-glucose pyrophosphorylase) (ADPGlc PPase).
 GN GLGC OR H11359.
 OS Haemophilus influenzae.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 CC Pasteurellaceae; Haemophilus.
 CC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=93550630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerkavane A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -!- CATALYTIC ACTIVITY: ATP + alpha-D-glucose 1-phosphate =
 CC diphosphate + ADP-glucose.
 CC -!- PATHWAY: Glycogen biosynthesis; first step.
 CC -!- SIMILARITY: Belongs to the bacterial/plant glucose-1-phosphate
 CC adenylyltransferase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U32815; AAC23006.1; ALT_INIT.
 CC PIR; B64119; B64119.
 CC TIGR; H11359; -;
 CC HAMAP; MF_00624; -; 1.
 CC InterPro; IPR005836; ADP_Glu_PPYP.
 CC InterPro; IPR005835; NTP_transferase.
 CC Pfam; PF00483; NTP_transferase; 1.
 CC PROSITE; PS00808; ADP_GLC_PYROPHOSPH_1; 1.
 CC PROSITE; PS00809; ADP_GLC_PYROPHOSPH_2; 1.
 CC PROSITE; PS00810; ADP_GLC_PYROPHOSPH_3; 1.
 CC Glycogen biosynthesis; Transferase; Nucleotidyltransferase;
 CC Complete proteome.
 CC KW SEQUENCE 437 AA; 49107 MW; E5E2284ACE5AD1D9 CRC64;
 CC
 CC Query Match 9.2%; Score 75.5; DB 1; Length 437;
 CC Best Local Similarity 26.7%; Pred. No. 5.1;
 CC Matches 27; Conservative 16; Mismatches 47; Indels 11; Gaps 3;
 CC
 CC QY 45 QRFYVYVMPKGYEYDQR---WILESGVGHYKLNKFSGRYLYVGTGFLTAGAAV 101
 CC Db 92 QERGFVDMFARQQIDSDTWYRGTDADAVYQNNAIKHNKHYRKYILLAGDHYKODYSV 151
 CC
 CC QY 102 REMDHFKFTADGTGKYDISSKANGHPRGRKNGKNGVMDKE 141
 CC Db 152 MLMMDHVNYSKACTVGCIEV-----PRSEAEHFGVWAYNE 185
 CC
 CC RESULT 7
 CC ALA3_HUMAN
 CC ID ALA3_HUMAN STANDARD; PRT; 1013 AA.
 CC AC P13637; Q16732; Q16735; Q969K5;
 CC DT 01-JAN-1990 (Rel. 13, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sodium/potassium-transferring ATPase alpha-3 chain (EC 3.6.3.9)
 DE (Sodium pump 3) (Na+/K+ ATPase 3) (Alpha (III)).
 GN ATP1A3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88255304; PubMed=2838329;
 RA Ovchinnikov Y.A., Monastyrskaya G.S., Broude N.E., Ushkaryov Y.A.,
 RA Melkov A.M., Smirnov Y.V., Malyshev I.V., Allikmets R.L.,
 RA Kostina M.B., Dulubova I.E., Kiyatkin N.I., Grishin A.V.,
 RA Modyanov N.N., Ovchinnikov Y.A.,
 RA Modyanov N.N., Sverdlov V.E.,
 RT "Family of human Na+, K+-ATPase genes. Structure of the gene for the
 RT catalytic subunit (alpha III-form) and its relationship with
 RT structural features of the protein.";
 RL FEBS Lett. 233:87-94 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Sverdlov E.D., Monastyrskaya G.S., Broude N.E., Ushkarev Y.A.,
 RA Melkov A.M., Smirnov Y.V., Malyshev I.V., Allikmets R.L.,
 RA Kostina M.B., Dulubova I.E., Kiyatkin N.I., Grishin A.V.,
 RA Modyanov N.N., Ovchinnikov Y.A.,
 RT "Family of human Na+(+)-ATPase genes. Structure of the gene of
 RT isoform alpha-III.";
 RL Dokl. Akad. Nauk SSSR 297:1488-1494 (1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan K., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McWan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Souffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [4]
 RP SEQUENCE OF 120-387; 494-538 AND 545-1013 FROM N.A.
 RX MEDLINE=87162481; PubMed=3030810;
 RA Ovchinnikov Y.A., Monastyrskaya G.S., Broude N.E., Allikmets R.L.,
 RA Ushkaryov Y.A., Melkov A.M., Smirnov Y.V., Malyshev I.V.,
 RA Dulubova I.E., Petrukhin K.E., Gryshin A.V., Sverdlov V.E.,
 RA Kiyatkin N.I., Kostina M.B., Modyanov N.N., Sverdlov E.D.,
 RT "The family of human Na+, K+-ATPase genes. A partial nucleotide
 RT sequence related to the alpha-subunit.";
 RL FEBS Lett. 213:73-80 (1987).
 RN [5]
 RP ERATUM.
 RA Ovchinnikov Y.A., Monastyrskaya G.S., Broude N.E., Allikmets R.L.,
 RA Ushkaryov Y.A., Melkov A.M., Smirnov Y.V., Malyshev I.V.,
 RA Dulubova I.E., Petrukhin K.E., Gryshin A.V., Sverdlov V.E.,
 RA Kiyatkin N.I., Kostina M.B., Modyanov N.N., Sverdlov E.D.,
 RL FEBS Lett. 214:375-375 (1987).
 RN [6]
 RP SEQUENCE OF 243-434 FROM N.A.

RX MEDLINE=87247232; PubMed=3036582;
 RA Sverdlov E.D., Monastyrskaya G.S., Broude N.E., Ushkaryov Y.A.,
 RA Allikmets R.L., Melkov A.M., Smirnov Y.V., Malyshev I.V.,
 RA Dulubova I.E., Petrukhin K.E., Grishin A.V., Kiyatkin N.I.,
 RA Kostina M.B., Sverdlov V.E., Modyanov N.N., Ovchinnikov Y.A.,
 RT "The family of human Na+, K+-ATPase genes. No less than five genes
 RT and/or pseudogenes related to the alpha-subunit.";
 RL FEBS Lett. 217:275-278 (1987).
 CC
 CC -!- FUNCTION: This is the catalytic component of the active enzyme,
 CC which catalyzes the hydrolysis of ATP coupled with the exchange of
 CC sodium and potassium ions across the plasma membrane. This action
 CC creates the electrochemical gradient of sodium and potassium ions,
 CC providing the energy for active transport of various nutrients.
 CC -!- CATALYTIC ACTIVITY: ATP + H2O + Na(+) (In) + K(+) (Out) = ADP +
 CC phosphate + Na(+) (Out) + K(+) (In).
 CC -!- SUBUNIT: Composed of three subunits: alpha (catalytic), beta and
 CC gamma.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
 CC ATPases). Subfamily IIC.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M37457; AAA51798.1; JOINED.
 DR EMBL; M37436; AAA51798.1; JOINED.
 DR EMBL; M37437; AAA51798.1; JOINED.
 DR EMBL; M37438; AAA51798.1; JOINED.
 DR EMBL; M37462; AAA51798.1; JOINED.
 DR EMBL; M37439; AAA51798.1; JOINED.
 DR EMBL; M37440; AAA51798.1; JOINED.
 DR EMBL; M37441; AAA51798.1; JOINED.
 DR EMBL; M37442; AAA51798.1; JOINED.
 DR EMBL; M37443; AAA51798.1; JOINED.
 DR EMBL; M37444; AAA51798.1; JOINED.
 DR EMBL; M37445; AAA51798.1; JOINED.
 DR EMBL; M37447; AAA51798.1; JOINED.
 DR EMBL; M37448; AAA51798.1; JOINED.
 DR EMBL; M37449; AAA51798.1; JOINED.
 DR EMBL; M37450; AAA51798.1; JOINED.
 DR EMBL; M37451; AAA51798.1; JOINED.
 DR EMBL; M37452; AAA51798.1; JOINED.
 DR EMBL; M37453; AAA51798.1; JOINED.
 DR EMBL; M37454; AAA51798.1; JOINED.
 DR EMBL; M37455; AAA51798.1; JOINED.
 DR EMBL; M37456; AAA51798.1; JOINED.
 DR EMBL; X12910; CAA31390.1; JOINED.
 DR EMBL; X12911; CAA31390.1; JOINED.
 DR EMBL; X12912; CAA31390.1; JOINED.
 DR EMBL; X12913; CAA31390.1; JOINED.
 DR EMBL; X12914; CAA31390.1; JOINED.
 DR EMBL; X12915; CAA31390.1; JOINED.
 DR EMBL; X12916; CAA31390.1; JOINED.
 DR EMBL; X12917; CAA31390.1; JOINED.
 DR EMBL; X12919; CAA31390.1; JOINED.
 DR EMBL; X12920; CAA31390.1; JOINED.
 DR EMBL; X12921; CAA31390.1; JOINED.
 DR EMBL; X12922; CAA31390.1; JOINED.
 DR EMBL; X12923; CAA31390.1; JOINED.
 DR EMBL; M28286; AAA52285.1; JOINED.
 DR EMBL; M28284; AAA52285.1; JOINED.
 DR EMBL; M28285; AAA52285.1; JOINED.
 DR EMBL; M28293; AAA52286.1; JOINED.
 DR EMBL; M28287; AAA52286.1; JOINED.
 DR EMBL; M35821; AAA52286.1; JOINED.
 DR EMBL; M35822; AAA52286.1; JOINED.
 DR EMBL; M28289; AAA52286.1; JOINED.
 DR EMBL; M28290; AAA52286.1; JOINED.

DR EMBL; M28291; AAA52286.1; JOINED.
 DR EMBL; M28292; AAA52286.1; JOINED.
 DR EMBL; M27577; AAA58380.1; --
 DR EMBL; M27570; AAA58380.1; JOINED.
 DR EMBL; M27573; AAA58380.1; JOINED.
 DR EMBL; BC009282; AAH09282.1; --
 DR EMBL; BC009394; AAH09394.1; --
 DR EMBL; BC015566; AAH15566.1; --
 DR PIR; S00801; S00801.
 DR HSSP; P04191; 1EUL.
 DR Genew; HGNC:801; ATP1A3.
 DR MIM; 182350; --
 DR GO; GO:0005890; C:sodium/potassium-exchanging ATPase complex; TAS.
 DR GO; GO:0005391; F:sodium/potassium-exchanging ATPase activity; TAS.
 DR GO; GO:0006810; P:transport; TAS.
 DR InterPro; IPR001757; ATPase E1-E2.
 DR InterPro; IPR006069; Cation ATPase.
 DR InterPro; IPR006068; Cation ATPase C.
 DR InterPro; IPR004014; Cation ATPase N.
 DR InterPro; IPR008250; E1-E2 ATPase_Reg.
 DR InterPro; IPR005834; Hydrolase.
 DR InterPro; IPR005775; Na/K ATPase alph.
 DR Pfam; PF00689; Cation ATPase C; 1.
 DR Pfam; PF00690; Cation ATPase N; 1.
 DR Pfam; PF00122; E1-E2 ATPase; 1.
 DR Pfam; PF00702; Hydrolase; 1.
 DR PRINTS; PR00119; CATATPASE.
 DR PRINTS; PR00121; NAKATPASE.
 DR TIGRPFAM; TIGR01106; ATPase-IIC X-K; 1.
 DR TIGRPFAM; TIGR01494; ATPase P-type; 5.
 DR PROSITE; PS00154; ATPase E1_E2; 1.
 DR Hydrolase; Sodium/potassium transport; Transmembrane; Phosphorylation;
 KW Magnesium; Metal-binding; ATP-binding; Multigene family.
 FT DOMAIN 1 77 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 78 98 POTENTIAL.
 FT DOMAIN 99 121 LUMENAL (POTENTIAL).
 FT TRANSMEM 122 142 POTENTIAL.
 FT DOMAIN 143 278 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 279 298 POTENTIAL.
 FT DOMAIN 299 310 LUMENAL (POTENTIAL).
 FT TRANSMEM 311 328 POTENTIAL.
 FT DOMAIN 329 762 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 763 782 POTENTIAL.
 FT DOMAIN 783 792 LUMENAL (POTENTIAL).
 FT TRANSMEM 793 813 POTENTIAL.
 FT DOMAIN 814 833 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 834 856 POTENTIAL.
 FT DOMAIN 857 908 LUMENAL (POTENTIAL).
 FT TRANSMEM 909 928 POTENTIAL.
 FT DOMAIN 929 941 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 942 960 POTENTIAL.
 Query Match 9.1%; Score 75; DB 1; Length 1013;
 Best Local Similarity 21.5%; Pred. No. 14;
 Matches 48; Conservative 23; Mismatches 56; Indels 96; Gaps 12;
 QY 13 AIFGAASDN-PIADR-----KCVISDGL-VNHER-KPGQFFP-----YVY 52
 DB 421 AVFKGGQDNIFVLKRVADGASALSKCIELSSGSKVLMREKVKVAELPFNSTKYQL 480
 QY 53 MIPKGTPEYDDQWILESVGGDH-----YKLNKFGRYL----- 86
 DB 481 SIHETEDPNDNRVLLVMKGAPEILDRCSILLQKGEQPLDEBWKAFQNAVYELGLGE 540
 QY 87 -YVGTEDYFL-----TAGAVREMDHFKTTADG----- 113
 DB 541 RVLGFCYHYLPESQFPKGFAP-DCDDVNFPTDMLCFVGLMSMTDPPRAAVPDVAVGKCRSA 599
 QY 114 -----TGKYDISSKANGPRSGRKNVGWVKMDGKHYFTVEN 149
 DB 600 GIKVIMVTGDHPITAKA-----IAKGVGLISEGNE---TVED 633

RESULT 8
 AIA3 RAT STANDARD; PRT; 1013 AA.
 AC P06887; Q61732; Q921G6;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sodium/potassium-transporting ATPase alpha-3 chain (EC 3.6.3.9)
 DE (Sodium pump 3) (Na+/K+ ATPase 3) (Alpha(III)).
 GN ATP1A3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87128908; PubMed=3028470;
 RA Shull G.E., Greeb J., Lingrel J.B.;
 RT "Molecular cloning of three distinct forms of the Na+,K+-ATPase
 RT alpha-subunit from rat brain.";
 RL Biochemistry 25:8125-8132(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Brain;
 RX MEDLINE=88032933; PubMed=2822682;
 RA Hara Y., Urayama O., Kawakami K., Nojima H., Nagamune H., Kojima T.,
 RA Ohta T., Nagano K., Nakao M.;
 RT "Primary structures of two types of alpha-subunit of rat brain
 RT Na+,K+-ATPase deduced from cDNA sequences.";
 RL J. Biochem. 102:43-58(1987).
 RN [3]
 RP SEQUENCE OF 1-930 FROM N.A.
 RX TISSUE=Brain, and Liver;
 RX MEDLINE=88033255; PubMed=2822726;
 RA Herrera V.L.M., Emanuel J.R., Ruiz-Opazo N., Levenson R.,
 RA Nadal-Ginard B.;
 RT "Three differentially expressed Na,K-ATPase alpha subunit isoforms:
 RT structural and functional implications.";
 RL J. Cell Biol. 105:1855-1865(1987).
 RN [4]
 RP SEQUENCE OF 243-434 FROM N.A.
 RX MEDLINE=87247232; PubMed=3036582;
 RA Sverdlov E.D., Monastyrskaya G.S., Broude N.E., Ushkaryov Y.A.,
 RA Allikmets R.L., Melkov A.M., Smirnov Y.V., Malyshev I.V.,
 RA Dulobova I.E., Petrukhin K.E., Grishin A.V., Kijatkin N.I.,
 RA Kosina M.B., Sverdlov V.E., Modyanov N.N., Ovchinnikov Y.A.;
 RT "The family of human Na+,K+-ATPase genes. No less than five genes
 RT and/or pseudogenes related to the alpha-subunit.";
 RL FEBS Lett. 217:275-278(1987).
 CC -!- FUNCTION: This is the catalytic component of the active enzyme,
 CC which catalyzes the hydrolysis of ATP coupled with the exchange of
 CC sodium and potassium ions across the plasma membrane. This action
 CC creates the electrochemical gradient of sodium and potassium ions,
 CC providing the energy for active transport of various nutrients.
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Na(+) (in) + K(+) (out) = ADP +
 CC phosphate + Na(+) (out) + K(+) (in).
 CC -!- SUBUNIT: Composed of three subunits: alpha (catalytic), beta and
 CC gamma.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type
 CC ATPases). Subfamily IIC.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M14513; AAA40777.1; --
 CC EMBL; X05883; CAA29307.1; --
 CC EMBL; M28648; AAA41672.1; --

Matches 26; Conservative 15; Mismatches 36; Indels 28; Gaps 5;

```
QY 39 HERKPGQFFVYVYVIMPKGTGTDQWILEVSGGHHYKKNKFSGRYLVYGTFFDYLTAQ 98
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 36 HEALTGTETQLTETTF---SKNYQDYEFLLDVIHG-----FQYFIYGTAAFFFLYLG 82
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 99 A-----AVREM-DHFKFTADGTGKYDIISSKANGHPRSRG 131
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 83 ALLLAEGPYTTCGAVRQIFGDYRTTICGKG---LSATVYGGPKGRG 124
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 12
F825 ARATH
ID F825 ARATH STANDARD; PRT; 164 AA.
AC O22711;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein At1g60740.
GN AT1g60740 OR F8A5.25.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
[1]
SEQUENCE FROM N.A.
STRAIN=cv. Columbia;
MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.H., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana";
RL Nature 408:816-820(2000).
CC -!- SIMILARITY: Belongs to the peroxiredoxin 2 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC
CC -----
CC EMBL; AC002292; AAB71961.1; -.
CC F825; G96632; G96632.
CC HSP; P30044; I4A0.
CC InterPro; IPR000866; AhpC-TSA.
CC Pfam; PF00578; AhpC-TSA; 1.
CC KW Hypothetical protein.
CC SQ SEQUENCE 164 AA; 17841 MW; 8D934C52C7311B84 CRC64;
```

Query Match 8.8%; Score 72.5; DB 1; Length 164;
Best Local Similarity 27.8%; Pred. No. 3.4;
Matches 25; Conservative 14; Mismatches 36; Indels 15; Gaps 4;

```
QY 72 GDHYKLNKSGRYLVYGTFFDYLTA---GAAREMDHFKTADGTGKY-----DISS 121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 GKAEELKSGIDELICFVNDPFWKAWGKTYQENKHVFKVADGSGBYTHLLGLELDK 122
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 122 KANGHPSRRGKNGVGMKDGKHYFTVENCQ 151
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 KGLG---IRSRFALLDLNLK--VTVANVE 147
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
ECAR_ECHCA
ID ECAR_ECHCA STANDARD; PRT; 616 AA.
AC Q90495;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ecarin precursor (EC 3.4.24.-).
OS Echin carinatus (Saw-scaled viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Viperinae; Echis.
OC NCBI_TaxID=40353;
[1]
SEQUENCE FROM N.A.; AND SEQUENCE OF 191-228; 353-357; 372-383;
RP 393-415; 446-453; 459-474; 553-561; 574-597 AND 610-616.
RC TISSUE=Venom gland;
RX MEDLINE=95151760; PubMed=7849037;
RA Nishida S., Fujita T., Kohno N., Atoda H., Morita T., Takeya H.,
RA Kido I., Paine M.J.I., Kawabata S.-I., Iwanaga S.;
RT "cDNA cloning and deduced amino acid sequence of prothrombin activator
RT (ecarin) from Kenyan Echis carinatus venom.";
RL Biochemistry 34:1771-1778(1995).
[2]
CHARACTERIZATION.
RC TISSUE=Venom;
RX MEDLINE=76260130; PubMed=956136;
RA Morita T., Iwanaga S., Suzuki T.;
RT "The mechanism of activation of bovine prothrombin by an activator
RT isolated from Echis carinatus venom and characterization of the new
RT active intermediates.";
RL J. Biochem. 79:1089-1108(1976).
CC -!- FUNCTION: Catalyzes the conversion of prothrombin to alpha-
CC thrombin through formation of a thrombin intermediate. Has
CC metalloprotease activity.
CC -!- COPACTOR: Binds 1 zinc ion and 1 calcium ion per subunit (By
CC similarity).
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC
CC -----
CC EMBL; D32212; BAA06910.1; -.
CC F825; A55796; A55796.
CC HSP; P18619; IFVL.
CC InterPro; IPR006586; ADAM cysteine.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC Pfam; PF00200; disintegrin; 1.
CC Pfam; PF01562; Pep M12B proper; 1.
CC Pfam; PF01421; Reptolysin; 1.
CC PRINTS; PR00289; DISINTEGRIN.
CC PRODOM; PD000664; DISINTEGRIN.
CC SMART; SM00608; ACR; 1.
CC SMART; SM00050; DISIN; 1.
CC PROSITE; PS00215; ADAM MERPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
```



```

DR HSP; P18708; INSF.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003599; AAA ATPase_cent.
DR InterPro; IPR003960; AAA_sub.
DR InterPro; IPR009010; ASP_decarb_fold.
DR InterPro; IPR003338; ATPaseVAT_N.
DR InterPro; IPR004201; Cdc48_2.
DR Pfam; PF00004; AAA; 2.
DR Pfam; PF02933; Cdc48_2; 1.
DR Pfam; PF02359; Cdc48_N; 1.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00674; AAA; 1.
KW Transport; Protein transport; Hydrolase; Endoplasmic reticulum;
KW Golgi stack; ATP-binding; Metal-binding; Magnesium; Repeat.
FT NP_BIND 257 264 ATP (POTENTIAL).
FT NP_BIND 537 544 ATP (POTENTIAL).
FT METAL 544 544 MAGNESIUM.
SQ SEQUENCE 742 AA; 81487 MW; 39611058DEAC4A32 CRC64;

Query Match 8.7%; Score 72; DB 1; Length 742;
Best Local Similarity 28.8%; Pred. No. 20;
Matches 34; Conservative 16; Mismatches 44; Indels 24; Gaps 7;

QY 8 FVLIALFGASNDPIA-----DRKIVISDGLVMHER-KPQGEPPYVY-----MIPKG 57
DB 54 FILSLGCHGSRDGNIALNALQRRHARVSTGDMVSFRVPPENFDLAMLLELEFYKKG 113
QY 58 TEYDQDWILSVGGDHVKLNKPSGRVLYVG---TFD-----YFLTAGAA-VREMDH 106
DB 114 TKSQVDAAALLST-----OLKRRKYNQVLTGVQKATPEYHGNTVILTVNRADVGGQDH 166

RESULT 15
SPR1 ARATH STANDARD; PRT; 303 AA.
AC Q22315; Q22314; Q39201;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pre-mRNA splicing factor SF2 (SR1 protein).
GN SF2 OR A1UG02840 OR F22D16.16 OR F22D16.30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95372342; PubMed=7644475;
RA Lazar G., Schaal T., Maniatis T., Goodman H.M.;
RT "Identification of a plant serine-arginine-rich protein similar to the
RL mammalian splicing factor SF2/ASF."
RL Proc. Natl. Acad. Sci. U.S.A. 92:7672-7676 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Su C.-L., Schuler M.A.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,

```

```

RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RL thaliana";
RL Nature 408:816-820 (2000).
CC -!- FUNCTION: Can promote splice site selection in vitro presumably by
CC antagonizing the effects of the Al heterogeneous nuclear
CC ribonucleoprotein. May have an essential function during early
CC plant development.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=SF2A;
CC IsoId=Q22315-1; Sequence=Displayed;
CC Name=2; Synonyms=SF2B;
CC IsoId=Q22315-2; Sequence=VSP_005859, VSP_005860;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- PTM: Extensively phosphorylated on serine residues in the RS
CC domain.
CC -!- SIMILARITY: Belongs to the splicing factor SR family.
CC -!- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC

```

```

EMBL; M98340; AAA32856.1; -
EMBL; AF001035; AAB71385.1; -
EMBL; AF001035; AAB71386.1; -
EMBL; AC009525; AAF02881.1; -
PIR; F86158; F86158.
PIR; S71185; S71185.
HSP; P19339; ISXL.
InterPro; IPR000504; RNA_rec_mot.
Pfam; PF00076; rrm; 2.
SMART; SM00360; RRM; 2.
PROSITE; PS00102; RRM; 2.
PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
KW mRNA processing; mRNA splicing; Nuclear protein; RNA-binding; Repeat;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 7 82 RNA-BINDING (RRM) 1.
FT DOMAIN 96 112 GLY-RICH (HINGE REGION).
FT DOMAIN 119 197 RNA-BINDING (RRM) 2.
FT DOMAIN 198 273 ARG/SER-RICH (RS DOMAIN).
FT DOMAIN 274 303 LYS/PRO/SER-RICH (PSK DOMAIN).
FT VARSPLIC 267 272 SRSRSR -> YGFTYD (in isoform 2).
FT VARSPLIC 273 303 Missing (in isoform 2).
FT CONFLICT 125 125 /FTID=VSP_005860.
FT CONFLICT 127 127 P -> W (IN REF. 1).
FT CONFLICT 139 140 MR -> IA (IN REF. 1).
SQ SEQUENCE 303 AA; 33729 MW; 6530F9CB628B8EFA CRC64;

```

```

Query Match 8.7%; Score 71.5; DB 1; Length 303;
Best Local Similarity 25.2%; Pred. No. 8.5;
Matches 26; Conservative 17; Mismatches 51; Indels 9; Gaps 2;

QY 35 DLVWHRKQGEPPYVYVMPKGE-----YDQDWILSVGGDHVKLNKPSGRVLY 87
DB 134 DLKDHWRKGDVCFQVYRDARTTGVVDTCTCYEDMKYALKL--DTEFRNFAFSNGYVR 191
QY 88 YGTFYFLTAGAAVREMDHFKFTADGTGKYDISSKANGHPRSR 130
DB 192 VREYDSKDRSPSRGSRGSRGSRGSRGSRGSRGSRGSR 234

```

Thu Sep 9 15:41:32 2004

us-10-768-874-2.rsp

Page 12

Search completed: September 9, 2004, 15:28:36
Job time : 26 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	825	100.0	152	5	077009	077009	simulium vi
2	88	10.7	712	2	085047	085047	oraxella c
3	84	10.2	868	10	07XMD8	07XMD8	coryza sativa
4	83.5	10.1	389	16	08A9R8	08A9R8	bacteroides
5	83.5	10.1	474	16	P74698	P74698	synchocystis
6	80.5	9.8	280	6	Q9TTR1	Q9TTR1	canis fami
7	80.5	9.8	329	6	Q9TTR2	Q9TTR2	canis fami
8	79.5	9.6	1048	2	Q9LBO9	Q9LBO9	arthrobacte
9	78	9.5	448	16	08DMZ4	08DMZ4	streptococ
10	77	9.3	501	2	Q3Z4J8	Q3Z4J8	streptococ
11	77	9.3	501	16	08DP07	08DP07	streptococ
12	77	9.3	557	16	Q97552	Q97552	clostridium
13	76.5	9.3	332	16	Q98RC7	Q98RC7	mycoplasma
14	76.5	9.3	411	2	Q8KKH3	Q8KKH3	bacillus ci
15	76.5	9.3	422	16	Q7V8P5	Q7V8P5	prochloroco
16	76.5	9.3	768	16	Q92J67	Q92J67	rickettsia

RESULT 2

Q85047 PRELIMINARY; PRT; 712 AA.
 ID Q85047
 AC Q85047
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transferrin binding protein B.
 GN TBPB.
 OS Moraxella catarrhalis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Moraxella.
 NCBI_TaxID=480;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=3;
 RX MEDLINE=98380363; PubMed=9712766;
 RA Myers L.E., Yang Y.P., Du R.P., Wang Q., Harkness R.E.,
 RA Schryvers A.B., Klein M.H., Loomore S.M.;
 RT "The transferrin binding protein B of moraxella catarrhalis elicits
 RT bactericidal antibodies and is a potential vaccine antigen."
 RL Infect. Immun. 66:4183-4192(1998).
 DR EMBL: AF039311; AAC34274.1;
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0004998; F:transferrin receptor activity; IEA.
 DR InterPro: IPR000437; Prok_lipoprot S.
 DR InterPro: IPR001677; Transferrin_bind.
 DR Pfam: PF01298; Lipoprotein_5; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 SQ SEQUENCE 712 AA; 76863 MW; C50808DD11AF2C09 CRC64;

Query Match 10.7%; Score 88; DB 2; Length 712;
 Best Local Similarity 27.8%; Pred. No. 3.2;
 Matches 37; Conservative 10; Mismatches 44; Indels 42; Gaps 6;
 QY 38 MHERKPGFFPYVYVIMPKGTEYDQWILESVGGDHKL-----KNKFSGR 84
 DB 276 WTDARKGQSFSSF-----ERRAGRYSAMSHYPSLLTDDKNKPNY 318
 QY 85 YLVYVG-----TDY---FLTAGAAREMDHFKFTADGTGKYDISSKANGHRSRGKNWGV 136
 DB 319 NDEYGHSSFTVDFSKSLTGLFLSLQDHKGKVKTKRYDINARIHGN-RFRGSATAI 377
 QY 137 MKDGE---KHFT 146
 DB 378 NKDNESKAKHPFT 390

RESULT 3

Q7XMD8 PRELIMINARY; PRT; 868 AA.
 ID Q7XMD8
 AC Q7XMD8
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE OSUNBA0061G20.21 protein.
 GN OSUNBA0061G20.21.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoideae; Oryza.
 NCBI_TaxID=4530;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
 RA Han B., Feng Q., Chen L., Fan D.L., Weng Q.J., Zhang L.,
 RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Liu Y., Li C., Li T.,
 RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
 RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
 RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
 RA Ren S.X., Lv G., Lin W., Gu W.G., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
 RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,

RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
 RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL662974; CA604665.1;
 SQ SEQUENCE 868 AA; 98333 MW; 28BD67842B4EB651 CRC64;
 Query Match 10.2%; Score 84; DB 10; Length 868;
 Best Local Similarity 24.8%; Pred. No. 11;
 Matches 33; Conservative 20; Mismatches 40; Indels 40; Gaps 7;
 QY 43 PGQFFPYVYVIMPKGTEYDQWIL-----ESVGGDH-----KLN 79
 DB 139 PAEAKNYTFTLPAGTCKVKRWILKKMAEQSFSGDLYQKYLKGTGTPNFTPFKLRD 198
 QY 80 KPSGRYLVGTFDYFLTAGAARE-----MDHFKFTADGTGKYD-----ISSKANGHP 127
 DB 199 HWD-EFVAYKTADGSLVFGDQIREAVRLTDAVEASSQGTFRDPRDELSALQTPPEH 257
 QY 128 -RSRGKN---MGV 136
 DB 258 GRTRGKGVIPWKI 270

RESULT 4

Q8A9R8 PRELIMINARY; PRT; 389 AA.
 ID Q8A9R8
 AC Q8A9R8
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative phosphatidylinositol-4-phosphate 5-kinase.
 GN BT0747.
 OS Bacteroides thetaiotaomicron.
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 NCBI_TaxID=818;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=24550858; PubMed=12663928;
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
 RA Chiang H.C., Hooper L.V., Gordon J.I.;
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."
 RL Science 299:2074-2076(2003).
 DR EMBL: AE016929; AA075854.1;
 DR GO: GO:0016301; P:kinase activity; IEA.
 DR InterPro: IPR003409; MORN.
 DR Pfam: PF02493; MORN; 14.
 KW Kinase; Complete proteome.
 SQ SEQUENCE 389 AA; 44241 MW; CE805074D37E851B CRC64;

Query Match 10.1%; Score 83.5; DB 16; Length 389;
 Best Local Similarity 24.4%; Pred. No. 4.2;
 Matches 33; Conservative 20; Mismatches 59; Indels 23; Gaps 7;
 QY 30 VISDGLVNHKPKGQEPFYVYVIMPKGTEYDQWILESVGGD---HYLKNKFSGRYL 86
 DB 72 VFKNEDVYEGEYIKKRGYGLYSPFDGKEYGQ-WFQDQOQHGKGIYFMNNRYDGMWF 130
 QY 87 V-----YGTFFYFLTAGAAREMDHFKFTADGTGKYDISSKANGH---PRSRGK---NW- 134
 DB 131 QDYQHAGTGM-YHNGDLYVGNWANDKREGEQTYTWANGAKYSCHWKNDKNGKGTMMND 189
 QY 135 -----GVMDGCKEH 143
 DB 190 DGCKYDGDWKKDDVRH 204

RESULT 5

P74698
 ID P74698
 AC P74698
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)

```
Db      7 MELNNILFVNTLLYGAASMKQAYFNKGTGELPCHFTNSQNISLDELVVFWQDQLVLVLY 66
Qy     40 ERKPGOEFPYYVMIPKG-TEYDDQRWILESGVDHYKLKNKFSGRY 85
Db      67 ELYRGKENPONVRKYKGRTSPDKONWTLR-----HNIQIKDGKGLY 108


RESULT 7
ID Q9TTF2 PRELIMINARY; PRT; 329 AA.
AC Q9TF2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B7-2 protein.
GN CD86
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
NCBI_TaxID=9615;
[1]
RP SEQUENCE FROM N.A.
EX MEDLINE=20033996; PubMed=10630300;
RA Yang S., Sim G.-K.;
RT "New Forms of Dog CD80 and CD86 Transcripts That Encode Secreted B7
RL Molecules.";
RL Immunogenetics 50:349-353(1999).
DR EMBL; AF106826; AAF17297.1; -.
DR HSSP; P33681; IDR9.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR IntraPro; IPR003596; IG_V.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 329 AA; 37774 MW; D98BB63437BF7B73 CRC64;

Query Match          9.8%; Score 80.5; DB 6; Length 329;
Best Local Similarity 27.1%; Pred. No. 6.9;
Matches 29; Conservative 13; Mismatches 38; Indels 27; Gaps

Qy    1 MSTQTGFVFVTILAIFQA-----SDNPIADRKCVISDGD-LVMH 39
Db    7 MELNNILFVNTLLYGAASMKQAYFNKGTGELPCHFTNSQNISLDELVVFWQDQLVLVLY 66

Qy   40 ERKPGOEFPYYVMIPKG-TEYDDQRWILESGVDHYKLKNKFSGRY 85
Db   67 ELYRGKENPONVRKYKGRTSPDKONWTLR-----HNIQIKDGKGLY 108


RESULT 8
Q9LBQ9 ID Q9LBQ9 PRELIMINARY; PRT; 1048 AA.
AC Q9LBQ9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glucodextranase.
GN GDA.
OS Arthroacter globiformis.
OC Bacteria; Actinobacteria; Actinomycetales; Actinomycetaceae;
OC Micrococineae; Micrococcaceae; Arthrobacter.
NCBI_TaxID=1665;
[1]
RP SEQUENCE FROM N.A.
RA STRAIN=i42;
RC Tomita R., Tonozuka T., Sakano Y.;
RT Molecular cloning of the Glucoctranase gene from Arthrobacter
RT globiformis i42, and Expression in Escherichia coli." ;
RL Submitted (Oct-1999) to the EMBL/Genbank/DDJB databases.
DL ENBL; AB033333; BAB20416.1; -.
DR GO; GO:0004339; F.glycan 1,4-alpha-metha-glucosidase activity; IEA.
DR GO; GO:0005976; P.polyaccharide metabolism; IEA.
```

```

DR InterPro: IPR006425; Glucan glucosid.
DR InterPro: IPR000165; Glyco_hydro.15.
DR InterPro: IPR008928; Glyco_trans.5np.
DR Pfam: PF00723; Glyco_hydro.15; 1.
DR TIGRfam: TIGR01535; Glucan glucosid; 1.
DR PROSITE: PS00820; GLUCOAMYLASE; 1.
SQ SEQUENCE 1048 AA; 109140 MW; 09D4D124B478DF4F CRC64;

Query Match          9.6%; Score 79.5; DB 2; Length 1048;
Best Local Similarity 23.9%; Pred. No. 40;
Matches 38; Conservative 16; Mismatches 60; Indels 45; Gaps 8;

QY 8 FVLTLAIFCAADN-----PIADKKCIIVISDGLVWHERKPGQFPYVYVIMPKGT 58
DB 396 PPTSEVDTGTCQNGIQIDETAFFILLANQIGRTDAGFYRELKPAAD-----YLVAGP 450
QY 59 EYDQDWILESVGGDHYKL-----KNKFSGRYLVYGTDFYFLTAGAA 100
DB 451 KTPQERW--EETGG--YSTSLASQIALAAADTAGKNGDAGSAVYRA-----TADFW 501
QY 101 VREMDHFETAD--GTGKYDISSKANGHPRSRG--KNWG 135
DB 502 QRSTKWMFTTNGPVGDGKYLRLISATGNPNDGATRDWG 540

RESULT 9
Q8DMZ4 PRELIMINARY; PRT; 448 AA.
AC Q8DMZ4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE Choline binding protein D.
GN CBPD OR SFR2006.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=171101;
RN [1] SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,
RA McLaren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,
RA Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rokey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
DR EMBL: AF008565; AAL00808.1; -.
DR PIR: C98122; C98122.
DR InterPro: IPR007921; CHAP.
DR InterPro: IPR002479; CW_binding.
DR InterPro: IPR003646; SH3_bac.
DR Pfam: PF05257; CHAP; 1.
DR Pfam: PF01473; CW_binding_1; 3.
DR SMART: SM00287; SH3b; 2.
DR PROSITE: PS00911; CHAP; 1.
KW Complete proteome.
SQ SEQUENCE 448 AA; 50396 MW; C136BE40CB530381 CRC64;

Query Match          9.5%; Score 78; DB 16; Length 448;
Best Local Similarity 23.8%; Pred. No. 19;
Matches 41; Conservative 19; Mismatches 38; Indels 74; Gaps 11;

QY 43 PGQEPFYVYVIMPKGTGYDDQKW-----ILSVGGDHY- 75
DB 217 PGEKVHYDQL-----EKDGKWLSTAYNGSVRYVQLEAVNKPNLGNVLSLTGGTHYF 271
QY 76 KLKPKFSGRYLVYGT--FDYFLTAGAA-----VREMDHFK---FTA-DGTGKY----- 117

```

```

DB 272 KIKSAIKTEPLVSATVIDYYPGEKVHYDQILEKDGKWLSTAYNGSVRYVQLEGVTS 331
QY 118 -----DISS-----KANG---HPRSRG--KNWGVKDGKHYF 145
DB 332 QNYQNSGNISSYSGNNSSTVGVKKINGSWYHFKSNGSKSTGWLKDGSSWY 383

RESULT 10
Q9Z4J8 PRELIMINARY; PRT; 501 AA.
AC Q9Z4J8;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE 1,4-beta-N-acetylmuramidase precursor (EC 3.2.1.17).
GN LYTC.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=M31;
RX MEDLINE=99340545; PubMed=10411730;
RA Garcia P., Gonzalez M.P., Garcia E., Garcia J.L., Lopez R.;
RT "The molecular characterization of the first autolytic lysozyme of
RT Streptococcus pneumoniae reveals evolutionary mobile domains.";
RL Mol. Microbiol. 33:128-138(1999).
DR EMBL: AJ009639; CA008765.1; -.
DR GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO: GO:0003796; F:lysozyme activity; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR002479; CW_binding.
DR Pfam: PF01473; CW_binding_1; 5.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 33
FT CHAIN 34 501 1,4-BETA-N-ACETYLMURAMIDASE.
SQ SEQUENCE 501 AA; 58682 MW; A6774B6DBF2EB704 CRC64;

Query Match          9.3%; Score 77; DB 2; Length 501;
Best Local Similarity 25.0%; Pred. No. 27;
Matches 31; Conservative 18; Mismatches 43; Indels 32; Gaps 8;

QY 34 GDLVHERKPGQEPFYVYVIMPKGTGYDDQKWILSVGGDHYKLR-----NKFSGRYL 86
DB 186 GMIQNEWLYDPAYSAYFYLSKDTG-YANQEW--QKVGKYYFKKNGYMARNEWQENY- 241
QY 87 VYGTDFYFLTAGAAVRE---MD--HFKFTADGTGKYDISSKANGHPRSRGKNWGVKGE 141
DB 242 -----YLTGGAMATDEVIMDGARYIFAASGELK-----EKKOLNVGV-VHRDCK 285
QY 142 KHYF 145
DB 286 RYFF 289

RESULT 11
Q8DP07 PRELIMINARY; PRT; 501 AA.
AC Q8DP07;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE 1,4-beta-N-acetylmuramidase (EC 3.2.1.17).
GN LYTC OR SP1431.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX Streptococcus.
OX NCBI_TaxID=171101;
RN [1] SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RX MEDLINE=21429245; PubMed=11544234;

```

RESULT 12	Q97G52	PRELIMINARY;	PRT;	557 AA.
ID	Q97G52			
AC	Q97G52;			
DT	01-OCT-2001 (TrEMBLrel. 18, Created)			
DT	01-OCT-2001 (TrEMBLrel. 18, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Extracellular neutral metalloprotease, NPPE.			
DE	CAC2517.			
GN	Clostridium acetobutylicum.			
OS	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1488;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 824 / DSM 792 / VKM B-1787;			
RC	MEDLINE=21359325; PubMed=11466286;			
RA	Noelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.			
RA	Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,			
RA	Tatsovs R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.S.			
RA	Bennett G.N., Koonin E.V., Smith D.R.;			
RT	"Genome sequence and comparative analysis of the solvent-producing			
RT	bacterium Clostridium acetobutylicum."			
RL	J. Bacteriol. 183:4823-4838 (2001).			
DR	EMBL; AB007750; AAK80471.1; -			
DR	PIR; A60152; A60152.			
DR	PIR; D97210; D97210.			
DR	MEROPS; M04.011; -			
DR	GO; GO:0005576; C:extracellular; IEA.			
DR	GO; GO:0004222; F:metalloendopeptidase activity; IEA.			
DR	GO; GO:0008270; P:zinc ion binding; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR001570; Peptidase M4.			
DR	InterPro; IPR005075; Peptidase M4 N.			
DR	InterPro; IPR006025; Pept M Zn BS.			

	RESULT 15			
ID	Q7V8P5	PRELIMINARY;	PRT;	422 AA.
AC	Q7V8B5;			
DT	01-OCT-2003	(TEMBLRrel. 25,	Created)	
DT	01-OCT-2003	(TEMBLRrel. 25,	Last sequence update)	
DT	01-OCT-2003	(TEMBLRrel. 25,	Last annotation update)	
DE		Page integrate.		
CN	PMT0290.			

Result No.	Query			ID	Description
	Score	Match	Length		
1	825	100.0	152	2	AAW74589
2	88	10.7	722	2	AAW43379
3	84.5	10.2	1297	2	ADC94925
4	80.5	9.8	280	2	ADY41078
5	80.5	9.8	329	2	AAW41076
6	79.5	9.6	1047	2	ABR35740
7	78	9.5	1129	6	ABRS2906
8	76.5	9.3	332	3	ADC23977
9	76.5	9.3	448	3	AAW71551
10	76	9.2	471	2	AAW61216
11	76	9.2	471	5	ABP54635
12	76	9.2	471	7	ADC45239
13	76	9.2	490	3	AAW81745
14	76	9.2	506	6	AAW20334
15	75.5	9.2	179	3	AAU38201
16	75.5	9.1	180	3	AAU10251
17	75	9.1	446	5	AAU76561
18	75	9.1	446	5	AAU76564
19	75	9.1	448	3	AAW71285
20	75	9.1	448	6	ABU02732
21	75	9.1	676	5	ABP29594
22	75	9.1	676	6	ABU46865
23	75	9.1	1013	7	ADB85134
24	75	9.1	1013	7	ADB55704
25	75	9.1	1013	7	ADB57722

CC can also be used for treating atherosclerosis of extremities, heart
 CC failure, hypertension, peripheral resistance, stenoses and in particular
 CC peripheral vasodilation. These proteins may also be used to suppress the
 CC immune system, to inhibit or prevent the development of antibodies or
 CC cellular immunity to a protein, treat graft rejection and autoimmune
 CC diseases. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-
 CC MAR-2003 to correct PI field.)

XX Sequence 152 AA;

Query Match 100.0%; Score 825; DB 2; Length 152;
 Best Local Similarity 100.0%; Pred. No. 8.1e-87;
 Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSITQFFVLTALFAGASDNPIADKCIIVISDGLVMSHERKPGQFFPYVYMIKGTET 60
 DB 1 MSITQFFVLTALFAGASDNPIADKCIIVISDGLVMSHERKPGQFFPYVYMIKGTET 60
 QY 61 DDQWILSVGGDHYKLNKFGRIYVGTFFYFLTAGAAREMDHFKFTADGTGKYDIS 120
 DB 61 DDQWILSVGGDHYKLNKFGRIYVGTFFYFLTAGAAREMDHFKFTADGTGKYDIS 120
 QY 121 SKANGHPRSRGKNWGMKDGKHYFTVENCQE 152
 DB 121 SKANGHPRSRGKNWGMKDGKHYFTVENCQE 152

RESULT 2

AY43379
 ID AAY43379 standard; protein; 722 AA.

XX AC AAY43379;

DT 26-JAN-2000 (first entry)

XX M. catarrhalis strain 3 tbp2 protein.

KW TbpB Gene; Tbp2; transferrin binding protein; diagnosis; otitis media;
 KW genetic immunisation; Moraxella infection; antigen; vaccine; detection;
 KW antitumour antibody production; therapy.

OS Moraxella catarrhalis.

XX WO9522947-A2.

XX 21-OCT-1999.

XX 12-APR-1999; 99WO-CA000307.

XX 14-APR-1998; 98US-00059584.

XX (CONN-) CONNAUGHT LAB LTD.

XX Myers LE, Schryvers AB, Harkness RE, Loesmore SM, Du R, Yang Y;
 PI Klein MH;

XX WPI; 1999-620376/53.

XX N-PSDB; AA231947.

PT Nucleic acid encoding transferrin binding protein 2 of Moraxella
 PT catarrhalis, useful for diagnostics, immunization and recombinant protein
 PT production.

XX Claim 7; Fig 4; 114pp; English.

XX This sequence is the Moraxella catarrhalis strain 3 transferrin binding
 CC protein (Tbp2) of the invention. The DNA sequence is also referred to as
 CC the TbpB gene. The TbpB gene is used to produce recombinant Tbp2; for
 CC identification or diagnosis of Moraxella, or for cloning related species,
 CC using hybridisation assays; and for genetic immunisation against
 CC Moraxella infections, e.g. otitis media. The Tbp2 proteins are useful as
 CC antigens, either in vaccines (including components of conjugate vaccines
 CC that contain antigens from other bacteria or from tumours, in which case

CC they elicit production of antitumour antibodies that may be coupled to
 CC chemotherapeutic agents or biologically active agents) or to raise
 CC antibodies (for use as diagnostic reagents and for treating Moraxella
 CC infections), also for detecting Moraxella antibodies

XX Sequence 722 AA;

Query Match 10.7%; Score 88; DB 2; Length 722;
 Best Local Similarity 27.8%; Pred. No. 0.81;
 Matches 37; Conservative 10; Mismatches 44; Indels 42; Gaps 6;

QY 38 MHERKPGQFFPYVYMIKGTETDQRMILSVGGDHYKL-----KNKFSGR 84
 DB 286 MTDKKGQSFSSP-----ERRADRYSAMSSHEVPSLITDDKNKPDNY 328
 QY 85 YLVYG-----TFDY---FLTAGAAREMDHFKFTADGTGKYDISKANGHPRSRGKNWGV 136
 DB 329 NDEYGHSEFTVDFSKSLTGLFENLQDHKGVTKTRYDINARIHGN-RFRGSATAI 387
 QY 137 MKDGF--KHVFT 146
 DB 388 MKDNESKAKHFT 400

RESULT 3

ADC94925
 ID ADC94925 standard; protein; 1297 AA.

XX AC ADC94925;

DT 01-JAN-2004 (first entry)

XX E. faecium protein sequence SEQ ID 4552.

KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.

OS Enterococcus faecium.

XX US6583275-B1.

XX 24-JUN-2003.

XX 30-JUN-1998; 98US-00107532.

XX 02-JUL-1997; 97US-0051571P.

XX 14-MAY-1998; 98US-0085598P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2003-799836/75.

XX N-PSDB; ADC91271.

PT New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.

XX Example 1; SEQ ID NO 4552; 243pp; English.

XX The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridising to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to a
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids is useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract

CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.
 XX
 SQ Sequence 1297 AA;
 Query Match 10.2%; Score 84.5; DB 7; Length 1297;
 Best Local Similarity 28.8%; Pred. No. 4.6;
 Matches 36; Conservative 14; Mismatches 36; Indels 39; Gaps 8;
 QY 30 VISDGLVMHE-----RKQGEFFYY-YMI-----PKGYDDQR--WIL-----ES 69
 DB 920 VIVDGEKVIHPICEPDSFVYKDSSETTYNAEYKIIESQNPGEKPEPDNDTWILRYDNA 979
 QY 70 VGGDHVYKLNKPSGRYL-----VYGTDFYELTAGAAVREMDHPKFTADGT 114
 DB 980 IGTIQLKQTGEENWQYLYTTTDEEDRLWVAFYATNFDY---REIRVKLDHEGHTLDGA 1036
 QY 115 GKVDI 119
 DB 1037 G-FDI 1040

RESULT 4
 ID AAY41078
 XX AAY41078 standard; protein; 280 AA.
 AC AAY41078;
 XX
 XX 20-DEC-1999 (first entry)
 DT
 XX Canine B7-2S protein.
 DE
 XX B7; CTLA4; T cell costimulatory protein; dog; cat; autoimmune disease;
 KW allergic reaction; infectious disease; tumor development; canine;
 KW graft rejection; inflammation; arthritis; atopic dermatitis.
 XX
 OS Canis familiaris.
 XX
 XX WO9947558-A2.
 XX
 XX 23-SEP-1999.
 XX
 XX 19-MAR-1999; 99WO-US006187.
 XX
 XX 19-MAR-1998; 98US-0078765P.
 XX 17-APR-1998; 98US-00062597.
 XX
 PA (HESK-) HESKA CORP.
 XX
 XX Sim G, Yang S, Sellins KS;
 XX
 XX WPI; 1999-571822/48.
 XX N-PSDB; AAZ27921, AAZ27923.
 XX
 XX New isolated B7 and CTLA4 nucleic acids, used to develop products for
 XX treating, e.g. autoimmune and atopic diseases.
 XX
 PS Claim 4; Page 109-111; 148pp; English.

XX
 XX The invention provides B7 and CTLA4 (T cell costimulatory proteins)
 XX encoding nucleic acid molecules from dogs and cats. The proteins can be
 XX expressed by standard recombinant methodology. The nucleic acid molecules
 XX and the encoded proteins can be used for preventing or treating diseases,
 XX e.g. autoimmune diseases, allergic reactions, infectious diseases, tumor
 XX development, graft rejection, inflammation, arthritis and atopic diseases
 XX such as atopic dermatitis. They can be used in mammals such humans, dogs,
 XX cats, cattle, sheep or pets. The products can also be used for detection,
 XX diagnosis and drug screening

XX
 SQ Sequence 280 AA;
 Query Match 9.8%; Score 80.5; DB 2; Length 280;
 Best Local Similarity 27.1%; Pred. No. 1.6;
 Matches 29; Conservative 13; Mismatches 38; Indels 27; Gaps 4;
 QY 1 MSITQSFFVLTALFAGAA-----SDNPIADRKCVISDGD-LVMH 39
 DB 7 MELNNILFVMTLLLYGAAMKSOAYFNKGTGELPCHFTNSQNSISLDELVVFWDQDKLVLY 66
 QY 40 ERKPGQEFFYYVYMPKG-TEYDDQRWILESVGGDHVYKLNKPSGRY 85
 DB 67 ELYRGKENPQVHRKYKGTSTFDKDNWTLRL-----HNIQIKDKGLY 108
 RESULT 5
 ID AAY41076
 XX AAY41076 standard; protein; 329 AA.
 AC AAY41076;
 XX
 XX 20-DEC-1999 (first entry)
 DT
 XX Canine B7-2 protein.
 DE
 XX B7; CTLA4; T cell costimulatory protein; dog; cat; autoimmune disease;
 KW allergic reaction; infectious disease; tumor development; canine;
 KW graft rejection; inflammation; arthritis; atopic dermatitis.
 XX
 OS Canis familiaris.
 XX
 XX WO9947558-A2.
 XX
 XX 23-SEP-1999.
 XX
 XX 19-MAR-1999; 99WO-US006187.
 XX
 XX 19-MAR-1998; 98US-0078765P.
 XX 17-APR-1998; 98US-00062597.
 XX
 PA (HESK-) HESKA CORP.
 XX
 XX Sim G, Yang S, Sellins KS;
 XX
 XX WPI; 1999-571822/48.
 XX N-PSDB; AAZ27913, AAZ27915.
 XX
 XX New isolated B7 and CTLA4 nucleic acids, used to develop products for
 XX treating, e.g. autoimmune and atopic diseases.
 XX
 PS Claim 4; Page 97-99; 148pp; English.
 XX
 XX The invention provides B7 and CTLA4 (T cell costimulatory proteins)
 XX encoding nucleic acid molecules from dogs and cats. The proteins can be
 XX expressed by standard recombinant methodology. The nucleic acid molecules
 XX and the encoded proteins can be used for preventing or treating diseases,
 XX e.g. autoimmune diseases, allergic reactions, infectious diseases, tumor
 XX development, graft rejection, inflammation, arthritis and atopic diseases
 XX such as atopic dermatitis. They can be used in mammals such humans, dogs,
 XX cats, cattle, sheep or pets. The products can also be used for detection,
 XX diagnosis and drug screening
 XX
 PS Claim 4; Page 329 AA;
 Query Match 9.8%; Score 80.5; DB 2; Length 329;
 Best Local Similarity 27.1%; Pred. No. 2;
 Matches 29; Conservative 13; Mismatches 38; Indels 27; Gaps 4;
 QY 1 MSITQSFFVLTALFAGAA-----SDNPIADRKCVISDGD-LVMH 39
 DB 7 MELNNILFVMTLLLYGAAMKSOAYFNKGTGELPCHFTNSQNSISLDELVVFWDQDKLVLY 66

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 15:29:25 ; Search time 128 Seconds
(without alignments)
380.819 Million cell updates/sec

Title: US-10-768-874-2
Perfect score: 825
Sequence: 1 MSITQSFVLTALFAGASD.....NWGMKGKHYFTVENCQE 152

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	825	100.0	152	US-10-218-699-2	Sequence 2, Appli
2	825	100.0	152	US-10-218-699-2	Sequence 2, Appli
3	825	100.0	152	US-10-218-699-2	Sequence 2, Appli
4	610	73.9	109	US-10-218-699-5	Sequence 5, Appli
5	610	73.9	109	US-10-218-699-5	Sequence 5, Appli
6	610	73.9	109	US-10-218-699-5	Sequence 5, Appli
7	130	15.8	22	US-10-218-699-3	Sequence 3, Appli
8	130	15.8	22	US-10-218-699-3	Sequence 3, Appli
9	130	15.8	22	US-10-218-699-3	Sequence 3, Appli
10	87	10.5	15	US-10-218-699-4	Sequence 4, Appli
11	87	10.5	15	US-10-218-699-4	Sequence 4, Appli
12	87	10.5	15	US-10-218-699-4	Sequence 4, Appli
13	80.5	9.8	280	US-10-768-874-4	Sequence 17, Appli
14	80.5	9.8	329	US-10-768-874-4	Sequence 17, Appli
15	78.5	9.5	721	US-10-437-963-193032	Sequence 193032,

16	78	9.5	1143	15	US-10-369-493-1499	Sequence 1499, Ap
17	76.5	9.3	332	12	US-10-440-503-244	Sequence 244, App
18	76.5	9.3	332	12	US-10-440-503-244	Sequence 244, App
19	76.5	9.3	332	14	US-10-146-772-244	Sequence 244, App
20	76.5	9.3	332	15	US-10-241-742-244	Sequence 244, App
21	76.5	9.3	332	15	US-10-440-523-244	Sequence 244, App
22	76	9.2	471	9	US-09-765-272-158	Sequence 158, App
23	76	9.2	490	10	US-09-769-744A-98	Sequence 98, Appli
24	75	9.1	424	16	US-10-229-148B-13	Sequence 13, Appli
25	75	9.1	446	9	US-09-870-472-3	Sequence 6, Appli
26	75	9.1	446	9	US-09-870-472-6	Sequence 6, Appli
27	75	9.1	676	12	US-10-282-122A-74789	Sequence 74789, A
28	75	9.1	983	16	US-10-408-765A-2112	Sequence 2112, Ap
29	75	9.1	1013	14	US-10-205-194-15	Sequence 15, Appli
30	75	9.1	1013	16	US-10-408-765A-379	Sequence 379, App
31	75	9.1	1013	16	US-10-408-765A-2113	Sequence 2113, App
32	75	9.1	1015	16	US-10-408-765A-206	Sequence 206, App
33	75	9.1	1015	16	US-10-408-765A-207	Sequence 207, App
34	74.5	9.0	878	12	US-10-282-122A-65010	Sequence 65010, A
35	74.5	9.0	1471	16	US-10-437-963-190303	Sequence 190303, A
36	73	8.8	572	14	US-10-156-761-10627	Sequence 10627, A
37	73	8.8	996	16	US-10-437-963-175079	Sequence 175079, A
38	73	8.8	1034	16	US-10-437-963-175053	Sequence 175053, A
39	72.5	8.8	343	12	US-10-282-122A-51283	Sequence 51283, A
40	72.5	8.8	428	12	US-10-282-122A-46513	Sequence 46513, A
41	72.5	8.8	911	14	US-10-043-344-107	Sequence 107, App
42	72	8.7	255	10	US-09-880-748-1849	Sequence 1849, App
43	72	8.7	255	12	US-10-293-418-1849	Sequence 1849, App
44	72	8.7	943	9	US-09-969-362-5	Sequence 5, Appli
45	71.5	8.7	358	12	US-10-282-122A-71100	Sequence 71100, A

ALIGNMENTS

RESULT 1
US-10-218-699-2
; Sequence 2, Application US/10218699
; Publication No. US20030012795A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; TITLE OF INVENTION: Salivary Gland of the Black Fly
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/10/218,699
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US/09/702,647
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 152
; TYPE: PPT
; ORGANISM: Simulium vittatum
US-10-218-699-2

Query Match	100.0%	Score 825;	DB 14;	Length 152;
Best Local Similarity	100.0%	Pred. No. 1.8e-86;		
Matches 152;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSITQSFVLTALFAGASDNPADRCIVISDGLVMHKKPGQEFFYVYVIMPKGTEY	60	
Db	1	MSITQSFVLTALFAGASDNPADRCIVISDGLVMHKKPGQEFFYVYVIMPKGTEY	60	
QY	61	DDQWILSVGGDHYKLNKFSGYLYCTDYFLTAGAAVREMDHFFKFTADGTGKYDIS	120	
Db	61	DDQWILSVGGDHYKLNKFSGYLYCTDYFLTAGAAVREMDHFFKFTADGTGKYDIS	120	
QY	121	SKANGHPRSRGKNWGMKDGKHYFTVENCQE	152	

Db 121 SKANGHPRSRGKNGWVKDGEKHYFTVENCQE 152

RESULT 2

US-10-288-740-2
; Sequence 2, Application US/10288740
; Publication No. US20030109447A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; Ribeiro, Jose M.C.
; Cupp, Eddie W.
; Swaim, Steven F.
; TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM
; SALIVARY GLAND OF THE BLACK FLY
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/288,740
FILING DATE: 06-NO. US20030109447A1-2002
CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/036,355A

FILING DATE: 06-MARCH-1998

ATTORNEY/AGENT INFORMATION:

NAME: Spruill, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: 5721-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919 420 2202

TELEFAX: 919 881 3175

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-288-740-2

Query Match 100.0%; Score 825; DB 14; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.8e-86;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSITQSFVLTALFCAASDNPIADRKCVISDGLVWHERKPGQFFYYVYVIMPKGTEY 60

Db 1 MSITQSFVLTALFCAASDNPIADRKCVISDGLVWHERKPGQFFYYVYVIMPKGTEY 60

QY 61 DQQRWILESGGDHYKLNKFGSRVLYVGTFFYFLTAGAAVREMDHFKFTADGTGKYDIS 120

Db 61 DQQRWILESGGDHYKLNKFGSRVLYVGTFFYFLTAGAAVREMDHFKFTADGTGKYDIS 120

QY 121 SKANGHPRSRGKNGWVKDGEKHYFTVENCQE 152

Db 121 SKANGHPRSRGKNGWVKDGEKHYFTVENCQE 152

RESULT 3

US-10-768-874-2
; Sequence 2, Application US/10768874
; Publication No. US20040142869A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.

; APPLICANT: Ribeiro, Jose M.C.
; Cupp, Eddie W.
; Swaim, Steven F.
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; Salivary Gland of the Black Fly
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/10/768,874
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US/10/218,699
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US/09/702,647
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Simulium vittatum
; ORGANISM: Simulium vittatum
US-10-768-874-2

Query Match 100.0%; Score 825; DB 16; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.8e-86;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSITQSFVLTALFCAASDNPIADRKCVISDGLVWHERKPGQFFYYVYVIMPKGTEY 60

Db 1 MSITQSFVLTALFCAASDNPIADRKCVISDGLVWHERKPGQFFYYVYVIMPKGTEY 60

QY 61 DQQRWILESGGDHYKLNKFGSRVLYVGTFFYFLTAGAAVREMDHFKFTADGTGKYDIS 120

Db 61 DQQRWILESGGDHYKLNKFGSRVLYVGTFFYFLTAGAAVREMDHFKFTADGTGKYDIS 120

QY 121 SKANGHPRSRGKNGWVKDGEKHYFTVENCQE 152

Db 121 SKANGHPRSRGKNGWVKDGEKHYFTVENCQE 152

RESULT 4

US-10-218-699-5
; Sequence 5, Application US/10218699
; Publication No. US20030012795A1
; GENERAL INFORMATION:

APPLICANT: Cupp, Mary S.

APPLICANT: Ribeiro, Jose M.C.

APPLICANT: Cupp, Eddie W.

APPLICANT: Swaim, Steven F.

TITLE OF INVENTION: Recombinant Vasoactive Protein from

Salivary Gland of the Black Fly

FILE REFERENCE: 35721/205407

CURRENT APPLICATION NUMBER: US/10/218,699

CURRENT FILING DATE: 2002-08-14

PRIOR APPLICATION NUMBER: US/09/702,647

PRIOR FILING DATE: 2000-10-31

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 109

TYPE: PRT

ORGANISM: Simulium vittatum

US-10-218-699-5

Query Match 73.9%; Score 610; DB 14; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.8e-62;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 KPGQFFYYVYVIMPKGTEYDQQRWILESGGDHYKLNKFGSRVLYVGTFFYFLTAGAAV 101

Db 1 KPGQFFYYVYVIMPKGTEYDQQRWILESGGDHYKLNKFGSRVLYVGTFFYFLTAGAAV 60

QY 102 REMDHFKFTADGTGKYDISSKANGHPRSRGKNGWVKDGEKHYFTVENC 150

Db 61 REMDHFKFTADGTGKYDISSKANGHPRSRGKNGWVKDGEKHYFTVENC 109

RESULT 5
US-10-288-740-5
; Sequence 5, Application US/10288740
; Publication No. US20030109447A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; ; Ribeiro, Jose M.C.
; ; Cupp, Eddie W.
; ; Swaim, Steven F.
; ;
; TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM
; ; SALIVARY GLAND OF THE BLACK FLY
; ;
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622
; ;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/288,740
; FILING DATE: 06-March-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,355A
; FILING DATE: 06-MARCH-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/POCKET NUMBER: 5721-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-288-740-5

Query Match 73.9%; Score 610; DB 14; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.8e-62;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 42 KPGQFFPVYVMIKGTEDDQWILSVGGDHYKLNKFKSGRYLVYGTDFYFLTAGAAV 101
Db 1 KPGQFFPVYVMIKGTEDDQWILSVGGDHYKLNKFKSGRYLVYGTDFYFLTAGAAV 60
Qy 102 REMDHFXTADGTGKYDISSKANGHPRSRGKNWGMKDGKHYFTVENC 150
Db 61 REMDHFXTADGTGKYDISSKANGHPRSRGKNWGMKDGKHYFTVENC 109

RESULT 6
US-10-768-874-5
; Sequence 5, Application US/10768874
; Publication No. US20040142869A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; ; Ribeiro, Jose M.C.
; ; Cupp, Eddie W.
; ; Swaim, Steven F.
; ;
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; ; SALIVARY GLAND OF THE BLACK FLY
; ;
; NUMBER OF SEQUENCES: 5

FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/10/768,874
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US/10/218,699
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US/09/702,647
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Simulium vittatum
US-10-768-874-5

Query Match 73.9%; Score 610; DB 16; Length 109;
Best Local Similarity 100.0%; Pred. No. 5.8e-62;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 42 KPGQFFPVYVMIKGTEDDQWILSVGGDHYKLNKFKSGRYLVYGTDFYFLTAGAAV 101
Db 1 KPGQFFPVYVMIKGTEDDQWILSVGGDHYKLNKFKSGRYLVYGTDFYFLTAGAAV 60
Qy 102 REMDHFXTADGTGKYDISSKANGHPRSRGKNWGMKDGKHYFTVENC 150
Db 61 REMDHFXTADGTGKYDISSKANGHPRSRGKNWGMKDGKHYFTVENC 109

RESULT 7
US-10-218-699-3
; Sequence 3, Application US/10218699
; Publication No. US20030012795A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; ; Ribeiro, Jose M.C.
; ; Cupp, Eddie W.
; ; Swaim, Steven F.
; ;
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; ; SALIVARY GLAND OF THE BLACK FLY
; ;
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/10/218,699
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US/09/702,647
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Simulium vittatum
US-10-218-699-3

Query Match 15.8%; Score 130; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.9e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 131 GKNWGMKDGKHYFTVENCQE 152
Db 1 GKNWGMKDGKHYFTVENCQE 22

RESULT 8
US-10-288-740-3
; Sequence 3, Application US/10288740
; Publication No. US20030109447A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; ; Ribeiro, Jose M.C.
; ; Cupp, Eddie W.
; ; Swaim, Steven F.
; ;
; TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM
; ; SALIVARY GLAND OF THE BLACK FLY
; ;
; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/288,740
; FILING DATE: 06-NO. US20030109447A1-2002
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US/09/036,355A
; FILING DATE: 06-MARCH-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5721-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2202
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-288-740-3

Query Match 15.8%; Score 130; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.9e-08;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 GKNWGVMDGKXHYFTVENCQ 152
DB 1 GKNWGVMDGKXHYFTVENCQ 22

RESULT 9
US-10-768-874-3
; Sequence 3, Application US/10768874
; Publication No. US20040142869A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; TITLE OF INVENTION: Salivary Gland of the Black Fly
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/10/768,874
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US/10/218,699
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US/09/702,647
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Simulium vittatum
US-10-768-874-3

Query Match 15.8%; Score 130; DB 16; Length 22;
Best Local Similarity 100.0%; Pred. No. 9.9e-08;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 131 GKNWGVMDGKXHYFTVENCQ 152
DB 1 GKNWGVMDGKXHYFTVENCQ 22

RESULT 10

US-10-218-699-4
; Sequence 4, Application US/10218699
; Publication No. US20030012795A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: Recombinant Vasoactive Protein from
; TITLE OF INVENTION: Salivary Gland of the Black Fly
; FILE REFERENCE: 35721/205407
; CURRENT APPLICATION NUMBER: US/10/218,699
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US/09/702,647
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Simulium vittatum
US-10-218-699-4

Query Match 10.5%; Score 87; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 KPGQFFPYVYVMPK 56
DB 1 KPGQFFPYVYVMPK 15

RESULT 11
US-10-288-740-4
; Sequence 4, Application US/10288740
; Publication No. US20030109447A1
; GENERAL INFORMATION:
; APPLICANT: Cupp, Mary S.
; APPLICANT: Ribeiro, Jose M.C.
; APPLICANT: Cupp, Eddie W.
; APPLICANT: Swaim, Steven F.
; TITLE OF INVENTION: RECOMBINANT VASOACTIVE PROTEIN FROM
; TITLE OF INVENTION: SALIVARY GLAND OF THE BLACK FLY
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave. Suite 310
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27622
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/288,740
; FILING DATE: 06-NO. US20030109447A1-2002
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US/09/036,355A
; FILING DATE: 06-MARCH-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray

REGISTRATION NUMBER: 32,943
REFERENCE/DOCKET NUMBER: 5721-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919 420 2202
TELEFAX: 919 881 3175
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-288-740-4

Query Match 10.5%; Score 87; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 KPGQEFPPYVYMIPK 56
Db 1 KPGQEFPPYVYMIPK 15

RESULT 12

US-10-768-874-4

Sequence 4, Application US/10768874
Publication No. US20040142869A1
GENERAL INFORMATION:
APPLICANT: Cupp, Mary S.
APPLICANT: Ribeiro, Jose M.C.
APPLICANT: Cupp, Eddie W.
APPLICANT: Swaim, Steven F.
TITLE OF INVENTION: Recombinant Vasoactive Protein from
TITLE OF INVENTION: Salivary Gland of the Black Fly
FILE REFERENCE: 35721/205407
CURRENT APPLICATION NUMBER: US/10/768,874
CURRENT FILING DATE: 2004-01-30
PRIOR APPLICATION NUMBER: US/10/218,699
PRIOR FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: US/09/702,647
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 15
TYPE: PRT
ORGANISM: Simulium vittatum
US-10-768-874-4

Query Match 10.5%; Score 87; DB 16; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 KPGQEFPPYVYMIPK 56
Db 1 KPGQEFPPYVYMIPK 15

RESULT 13

US-10-790-396-17

Sequence 17, Application US/10790396
Publication No. US20040157296A1
GENERAL INFORMATION:
APPLICANT: Sim, Gek-Kee
APPLICANT: Yang, Shumin
APPLICANT: Sellins, Karen S.
TITLE OF INVENTION: NOVEL FORMS OF T CELL COSTIMULATORY
TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
FILE REFERENCE: IM-1-Cl-1-CT
CURRENT APPLICATION NUMBER: US/10/790,396
CURRENT FILING DATE: 2004-03-01
PRIOR APPLICATION NUMBER: US/09/646,561

PROTEINS, NUCLEIC

PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/078,765
PRIOR FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: 09/062,597
PRIOR FILING DATE: 1998-04-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 17
LENGTH: 280
TYPE: PRT
ORGANISM: Canis familiaris
US-10-790-396-17

Query Match 9.8%; Score 80.5; DB 16; Length 280;
Best Local Similarity 27.1%; Pred. No. 1.3;
Matches 29; Conservative 13; Mismatches 38; Indels 27; Gaps 4;

Qy 1 MSITQSPFVLTLAIFGAA-----SDNPIADRKCVISDGD-LVMH 39
Db 7 MELANNILFVMTLLYGAASMSQAYFNKTGELPCHFTNSQNISLDELVVFWDQDKLVLY 66
Qy 40 ERKPGQEFPPYVYMIPKG-TEYDDQRMWILESGDGHYKLNKPKSGRY 85
Db 67 ELYRGKENPQNVHRKYGRTSFDKNWTLRL-----HNIQIKDKGLY 108

RESULT 14

US-10-790-396-7

Sequence 7, Application US/10790396
Publication No. US20040157296A1
GENERAL INFORMATION:
APPLICANT: Sim, Gek-Kee
APPLICANT: Yang, Shumin
APPLICANT: Sellins, Karen S.
TITLE OF INVENTION: NOVEL FORMS OF T CELL COSTIMULATORY
TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
FILE REFERENCE: IM-1-Cl-1-CT
CURRENT APPLICATION NUMBER: US/10/790,396
CURRENT FILING DATE: 2004-03-01
PRIOR APPLICATION NUMBER: US/09/646,561
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 60/078,765
PRIOR FILING DATE: 1998-03-19
PRIOR APPLICATION NUMBER: 09/062,597
PRIOR FILING DATE: 1998-04-17
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 329
TYPE: PRT
ORGANISM: Canis familiaris
US-10-790-396-7

Query Match 9.8%; Score 80.5; DB 16; Length 329;
Best Local Similarity 27.1%; Pred. No. 1.6;
Matches 29; Conservative 13; Mismatches 38; Indels 27; Gaps 4;

Qy 1 MSITQSPFVLTLAIFGAA-----SDNPIADRKCVISDGD-LVMH 39
Db 7 MELANNILFVMTLLYGAASMSQAYFNKTGELPCHFTNSQNISLDELVVFWDQDKLVLY 66
Qy 40 ERKPGQEFPPYVYMIPKG-TEYDDQRMWILESGDGHYKLNKPKSGRY 85
Db 67 ELYRGKENPQNVHRKYGRTSFDKNWTLRL-----HNIQIKDKGLY 108

RESULT 15

US-10-437-963-193032

Sequence 193032, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.

